

GenScore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

1.0M protein - protein search, using SW model

Run on: January 1, 2001, 21:57:25 ; Search time 82.04 Seconds

(without alignments)
162,967 Million cell updates/sec

Database: 2147

1 MAHRRQELADYVVALHDD.....CPHVRKAWI-AMLIQILVVS 391

Scoring table:

BLAST2
Gapop: 10.0 , Gapext: 0.5

Searched: 268185 seqs, 44194795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database:

1: /S10S6/qqdata/geneseq/AA1980.DAT*
2: /S10S6/qqdata/geneseq/AA1981.DAT*
3: /S10S6/qqdata/geneseq/AA1982.DAT*
4: /S10S6/qqdata/geneseq/AA1983.DAT*
5: /S10S6/qqdata/geneseq/AA1984.DAT*
6: /S10S6/qqdata/geneseq/AA1985.DAT*
7: /S10S6/qqdata/geneseq/AA1986.DAT*
8: /S10S6/qqdata/geneseq/AA1987.DAT*
9: /S10S6/qqdata/geneseq/AA1988.DAT*
10: /S10S6/qqdata/geneseq/AA1989.DAT*
11: /S10S6/qqdata/geneseq/AA1990.DAT*
12: /S10S6/qqdata/geneseq/AA1991.DAT*
13: /S10S6/qqdata/geneseq/AA1992.DAT*
14: /S10S6/qqdata/geneseq/AA1993.DAT*
15: /S10S6/qqdata/geneseq/AA1994.DAT*
16: /S10S6/qqdata/geneseq/AA1995.DAT*
17: /S10S6/qqdata/geneseq/AA1996.DAT*
18: /S10S6/qqdata/geneseq/AA1997.DAT*
19: /S10S6/qqdata/geneseq/AA1998.DAT*
20: /S10S6/qqdata/geneseq/AA1999.DAT*
21: /S10S6/qqdata/geneseq/AA2000.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length DB	ID	Description
1	694	42.4	402	Porcine acylglucosaminase
2	691	42.2	16	Porcine acylglucosaminase
3	668	41.1	16	Porcine acylglucosaminase
4	645	40.0	16	Porcine acylglucosaminase
5	94	4.4	532	Alkaline Bacillus
6	93	4.3	688	Corn sulphate prot
7	92.5	4.3	659	Streptococcus pneumoniae
8	92	4.3	368	Streptococcus pneumoniae
9	90.5	4.2	678	Streptococcus pneumoniae
10	88	4.1	483	Alpha-amylase vari
11	88	4.1	999	Hovine lysosomal a
12	87.5	4.1	999	Hovine lysosomal a

13	87.5	4.1	1161	878169
14	87.5	4.1	1161	W24062
15	87.5	4.1	1161	W24824
16	87.5	4.1	1161	W65104
17	87.5	4.1	1161	W60004
18	87.5	4.1	1161	W73445
19	87	4.1	999	W26682
20	86	4.0	483	W12139
21	86	4.0	485	W18156
22	86	4.0	485	W15100
23	86	4.0	485	W48261
24	86	4.0	485	W25151
25	86	4.0	485	W15416
26	86	4.0	485	W15422
27	86	4.0	485	W07382
28	86	4.0	485	W07392
29	85	4.0	483	W12140
30	84.5	3.9	216	W06232
31	84.5	3.9	580	W45883
32	84.5	3.9	580	W01742
33	84.5	3.9	580	W11707
34	84.5	3.9	580	W44914
35	84.5	3.9	729	W67804
36	84.5	3.9	729	W24631
37	84.5	3.9	741	W24631
38	84.5	3.9	750	W47806
39	84	3.9	483	W12137
40	84	3.9	483	W12138
41	84	3.9	483	W14444
42	84	3.9	591	W55099
43	82.5	3.8	268	W21746
44	82	3.8	405	W85949
45	82	3.8	559	W47113

ALIGNMENTS

RESULT 1	
ID	R79931
AC	R79931 standard: Protein: 402 AA.
XX	
AC	R79931:
XX	
DT	09-MAY-1996 (first entry)
XX	
DE	Porcine acylglucosaminase-2-epimerase mutant.
XX	
KW	Porcine acylglucosaminase-2-epimerase; N-acetylglucosamine
KW	N-acetylglucosaminic acid; trans-acting; enzymatic production;
KW	mutant.
XX	
OS	Sus scrofa.
XX	
FT	Key
FT	Misc-difference 149
FT	/note- "wild type Asp subst. with Glu"
FT	Misc-difference 269
FT	/note- "wild type His subst. with Tyr"
FT	Misc-difference 417
FT	/note "wild type Ser subst. with Arg"
FT	Misc-difference 418
FT	/note- "wild type Glu subst. with Ala"
XX	
XX	W09526399-AL.
XX	
XX	05-OCT-1995.
XX	
XX	24-MAR-1995: 95W07-TP00541.
XX	
XX	09-SEP-1994: 94JP-0276344.
XX	
XX	25-MAR-1994: 94JP-0056271.
XX	

PA (MAKU) MAKUKIN SHOYU KK.
 XX
 P1 Maru I, Ohia Y, Tsukada Y?
 XX
 DR WP1: 1995-451420/45.
 XX
 PT Recombinant acylglucosamine-2-epimerase with reatin binding activity
 PT used in enzyme production of N-acetylmannosamine and
 PT N-acetylneuraminate acid
 XX
 PS claim 13: Pages 63-64; 74pp; Japanese.
 XX
 CC R79941 is a porcine acylglucosamine-2-epimerase (AZP) mutant,
 CC with reatin binding activity. AZP can be used for the enzymatic
 CC produ. of N-acetylmannosamine and N-acetylneuraminate acid.
 XX
 SO Sequence 402 AA;
 Query Match 42.48; Score 694; DB 16; Length 402;
 Best Local Similarity 44.88; Pred. No. 80-85;
 Matches 144; Conservative 77; Mismatches 154; Indels 48; Gaps 8;
 QY 5 KRGELAGUYVALHQD---VLPFRKYSIDROGGYFTCLLRKQAGVDTTKETMLQNKQV 61
 1
 DB 4 KRELIDAKKGVVPELDVMALEWLESHDEBQITTELDQIDQVYDILKYVWLPQRY 62
 1
 QY 62 WQFAVFNKLE--PKQWLEIARIAGVFLAKGR--DQGNWYFALDQEKRLQRYNVF 117
 1
 DB 63 WMYETLYKLTQTHPQLIDAKKQGETLITHTVAPKKCATVLTIDTPKQVQSIT 122
 1
 QY 118 STGFAMARFSQYALASAGQEKATALQAVYNNVLK KQINRKQYEKSYPTTRPKSLAVP 176
 1
 DB 123 SEETYLAMMELWETVATYSEAVQMDQIVHWETDPSQIDRPIIPAVASESMVPP 182
 1
 QY 177 MLLANLTLEM-----EWLLPTVEEVLAVTVKVMIDELDELDELGRVA 221
 1
 DB 183 MMLTLEQVLEQDEELQRYAQLQHW-----GARRILIQVQDQV-----VLEN 228
 1
 QY 222 VTPRGFVDSFEKELNLNGRGTEAMWFMMDIAQKSSDRQLQEQAT-AVNLNTELEVAMDEE 280
 1
 DB 229 VSDQGEISQGLGHPGPALEAGWLTITHSISQDAKIDAVDITLITPISQWAD 288
 1
 QY 281 FGGTFEFLDKQIHPPQOLEMQLMWVILETLVALAKSHQATQDEKQWQFERVHHYAMS 340
 1
 DB 289 IQGILYGLDADQLPQLQLEWAMKLPWPHSEAMIALIMYSESQDPAITLITYVAEYIT 348
 1
 QY 341 HFAIDFEGEMFGYLNRRGQVLLNKKGKMGCTIVPVALMIGCATLQDPVS 391
 1
 DB 349 QIDQYQEWLYIDREKVALILKQPKIKGLHPVPELAMEVMSALLS 399
 1
 RESULT 2
 ID R79928 standard; Porcine; 402 AA.
 AC R79928;
 XX
 DT 09 MAY 1996 (first entry)
 XX
 DE Porcine acylglucosamine-2-epimerase.
 XX
 KW Porcine acylglucosamine-2-epimerase; N-acetylmannosamine;
 KW N-acetylneuraminate acid; reatin-binding; enzymatic production.
 XX
 OS Suis serofa.
 XX
 PN W09526499 AL.
 XX
 PD 05 OCT 1995.
 XX
 PP 24 MAR 1995; 9-860-1000541.
 XX

PR 09 SEP 1994; 94JP-0216343.
 PR 25-MAR-1994; 94JP-0056271.
 XX
 PA (MAKU) MAKUKIN SHOYU KK.
 XX
 P1 Maru I, Ohia Y, Tsukada Y?
 XX
 DR WP1: 1995-451420/45.
 XX
 PT Recombinant acylglucosamine-2-epimerase with reatin binding activity
 PT used in enzyme production of N-acetylmannosamine and
 PT N-acetylneuraminate acid
 XX
 PS claim 2: Pages 53-54; 74pp; Japanese.
 XX
 CC R79928 porcine acylglucosamine-2-epimerase (AZP),
 CC with reatin binding activity. AZP can be used for the enzymatic
 CC produ. of N-acetylmannosamine and N-acetylneuraminate acid.
 XX
 SO Sequence 402 AA;
 Query Match 42.28; Score 691; DB 16; Length 402;
 Best Local Similarity 45.08; Pred. No. 1-76-64;
 Matches 144; Conservative 74; Mismatches 155; Indels 48; Gaps 8;
 QY 5 KRGELAGUYVALHQD---VLPFRKYSIDROGGYFTCLLRKQAGVDTTKETMLQNKQV 61
 1
 DB 4 KRELIDAKKGVVPELDVMALEWLESHDEBQITTELDQIDQVYDILKYVWLPQRY 62
 1
 QY 62 WQFAVFNKLE--PKQWLEIARIAGVFLAKGR--DQGNWYFALDQEKRLQRYNVF 117
 1
 DB 63 WMYETLYKLTQTHPQLIDAKKQGETLITHTVAPKKCATVLTIDTPKQVQSIT 122
 1
 QY 118 STGFAMARFSQYALASAGQEKATALQAVYNNVLK KQINRKQYEKSYPTTRPKSLAVP 176
 1
 DB 123 SEETYLAMMELWETVATYSEAVQMDQIVHWETDPSQIDRPIIPAVASESMVPP 182
 1
 QY 177 MLLANLTLEM-----EWLLPTVEEVLAVTVKVMIDELDELDELGRVA 221
 1
 DB 183 MMLTLEQVLEQDEELQRYAQLQHW-----GARRILIQVQDQV-----VLEN 228
 1
 QY 222 VTPRGFVDSFEKELNLNGRGTEAMWFMMDIAQKSSDRQLQEQAT-AVNLNTELEVAMDEE 280
 1
 DB 229 VSDQGEISQGLGHPGPALEAGWLTITHSISQDAKIDAVDITLITPISQWAD 288
 1
 QY 281 FGGTFEFLDKQIHPPQOLEMQLMWVILETLVALAKSHQATQDEKQWQFERVHHYAMS 340
 1
 DB 289 IQGILYGLDADQLPQLQLEWAMKLPWPHSEAMIALIMYSESQDPAITLITYVAEYIT 348
 1
 QY 341 HFAIDFEGEMFGYLNRRGQVLLNKKGKMGCTIVPVALMIGCATLQDPVS 391
 1
 DB 349 QIDQYQEWLYIDREKVALILKQPKIKGLHPVPELAMEVMSALLS 399
 1
 RESULT 3
 ID R79940 standard; Porcine; 419 AA.
 AC R79940;
 XX
 DT 09-MAY-1996 (first entry)
 XX
 DE Porcine acylglucosamine-2-epimerase mutant.
 XX
 KW Porcine acylglucosamine-2-epimerase; N-acetylmannosamine;
 KW N-acetylneuraminate acid; reatin binding; enzymatic production;
 KW mutant.
 XX
 OS Suis serofa.
 XX
 PN Key Location/Qualifiers
 XX

F1	Misc-difference	10	/note	"wild type	Ala	subst.	with Val"
F1	Misc-difference	21	/note	"wild type	Arg	subst.	with Ser"
F1	Misc-difference	23	/note	"wild type	Met	subst.	with Ile"
F1	Misc-difference	27	/note	"wild type	Leu	subst.	with Met"
F1	Misc-difference	34	/note	"wild type	Arg	subst.	with Glu"
F1	Misc-difference	47	/note	"wild type	Arg	subst.	with Glu"
F1	Misc-difference	51	/note	"wild type	Asp	subst.	with His"
F1	Misc-difference	71	/note	"wild type	Lys	subst.	with Thr"
F1	Misc-difference	72	/note	"wild type	Leu	subst.	with Phe"
F1	Misc-difference	75	/note	"wild type	His	subst.	with Arg"
F1	Misc-difference	78	/note	"wild type	Pro	subst.	with Val"
F1	Misc-difference	93	/note	"wild type	Arg	subst.	with Ser"
F1	Misc-difference	94	/note	"wild type	His	subst.	with Tyr"
F1	Misc-difference	101	/note	"wild type	Gly	subst.	with Gly"
F1	Misc-difference	110	/note	"wild type	Arg	subst.	with Glu"
F1	Misc-difference	120	/note	"wild type	Ser	subst.	with Thr"
F1	Misc-difference	136	/note	"wild type	Arg	subst.	with Lys"
F1	Misc-difference	139	/note	"wild type	Ala	subst.	with Gly"
F1	Misc-difference	141	/note	"wild type	Ala	subst.	with Met"
F1	Misc-difference	142	/note	"wild type	Ala	subst.	with Met"
F1	Misc-difference	145	/note	"wild type	Arg	subst.	with His"
F1	Misc-difference	149	/note	"wild type	Ser	subst.	with Arg"
F1	Misc-difference	155	/note	"wild type	Val	subst.	with Ile"
F1	Misc-difference	163	/note	"wild type	Ser	subst.	with Ala"
F1	Misc-difference	171	/note	"wild type	Pro	subst.	with Ser"
F1	Misc-difference	173	/note	"wild type	Ala	subst.	with Thr"
F1	Misc-difference	174	/note	"wild type	Val	subst.	with Leu"
F1	Misc-difference	176	/note	"wild type	Ser	subst.	with Thr"
F1	Misc-difference	178	/note	"wild type	Ser	subst.	with Pro"
F1	Misc-difference	187	/note	"wild type	Cys	subst.	with Asn"
F1	Misc-difference	199	/note	"wild type	Leu	subst.	with Met"
F1	Misc-difference	200	/note	"wild type	Ala	subst.	with Thr"
F1	Misc-difference	201	/note	"wild type	Gly	subst.	with Asp"
F1	Misc-difference	202	/note	"wild type	Arg	subst.	with Lys"
F1	Misc-difference	208	/note	"wild type	His	subst.	with Asp"
F1	Misc-difference	212	/note	"wild type	Arg	subst.	with His"
F1	Misc-difference	224	/note	"wild type	Arg	subst.	with His"

PI	Misc-difference	234	/note-	"wild type Ala substd. with Val"
ET	Misc-difference	237	/note-	"wild type gln substd. with lys"
ET	Misc-difference	249	/note-	"wild type Ser substd. with Pro"
ET	Misc-difference	258	/note	"wild type Ala substd. with Thr"
ET	Misc-difference	259	/note-	"wild type Arg substd. with gln"
ET	Misc-difference	260	/note-	"wild type His substd. with Tyr"
ET	Misc-difference	261	/note-	"wild type Ser substd. with Ala"
ET	Misc-difference	263	/note-	"wild type Ser substd. with Leu"
ET	Misc-difference	266	/note-	"wild type Ser substd. with Lys"
ET	Misc-difference	269	/note	"wild type Ala substd. with pro"
ET	Misc-difference	270	/note-	"wild type Arg substd. with gln"
ET	Misc-difference	272	/note-	"wild type Ala substd. with Arg"
ET	Misc-difference	275	/note	"wild type Val substd. with Ile"
ET	Misc-difference	282	/note-	"wild type Thr substd. with Lys"
ET	Misc-difference	287	/note-	"wild type Arg substd. with His"
ET	Misc-difference	288	/note-	"wild type Ala substd. with pro"
ET	Misc-difference	300	/note-	"wild type Asp substd. with gln"
ET	Misc-difference	317	/note-	"wild type Gly substd. with Asp"
ET	Misc-difference	328	/note-	"wild type Ser substd. with Thr"
ET	Misc-difference	329	/note-	"wild type Ser substd. with Arg"
ET	Misc-difference	343	/note-	"wild type Gln substd. with Asp"
ET	Misc-difference	348	/note-	"wild type Arg substd. with Asn"
ET	Misc-difference	364	/note-	"wild type Arg substd. with His"
ET	Misc-difference	393	/note-	"wild type Arg substd. with gln"
ET	Misc-difference	395	/note-	"wild type Met substd. with Ile"
ET	Misc-difference	399	/note-	"wild type Ser substd. with Gly"
ET	Misc-difference	402	/note-	"wild type Ser substd. with Gln"
ET	Misc-difference	403..419	/note-	"wild type Ala substd. with Gly"
ET	Misc-difference	403..419	/note-	"C-terminal addition to the wild type sequence"
XX	W09526399-A1.			
XX	05-oct-1995.			
XX	24-MAR-1995;	95WO-JP00541.		
XX	09-SEP-1994;	94JP-0216333.		
XX	25-MAR-1994;	94JP-0056271.		
PA	(MARU) MARUKIN SHOYU KK.			
PI	Maru I, Ohta Y, Tsukada Y;			
XX	WP1; 1995-351320/45.			

[illegible]

F0	Misc-difference	72	/noct-	"wild type Lys subst., with Thr"
F1	Misc-difference	76	/noct-	"wild type Leu subst., with Phe"
F2	Misc-difference	77	/noct-	"wild type His subst., with Arg"
F3	Misc-difference	78	/noct-	"wild type Arg subst., with His"
F4	Misc-difference	79	/noct-	"wild type Pro subst., with Ala"
F5	Misc-difference	94	/noct-	"wild type Gln subst., with Glu"
F6	Misc-difference	101	/noct-	"wild type His subst., with Tyr"
F7	Misc-difference	120	/noct-	"wild type Gln subst., with Gly"
F8	Misc-difference	137	/noct-	"wild type Ser subst., with Thr"
F9	Misc-difference	149	/noct-	"wild type Val subst., with Ala"
F10	Misc-difference	149	/noct-	"wild type Ala subst., with Gly"
F11	Misc-difference	141	/noct-	"wild type Ala subst., with Val"
F12	Misc-difference	145	/noct-	"wild type Asp subst., with Trp"
F13	Misc-difference	159	/noct-	"wild type Asp subst. I with Glu"
F14	Misc-difference	162	/noct-	"wild type Arg subst., with Glu"
F15	Misc-difference	171	/noct-	"wild type Pro subst., with Ala"
F16	Misc-difference	174	/noct-	"wild type Pro subst., with Gln"
F17	Misc-difference	176	/noct-	"wild type Val subst., with Pro"
F18	Misc-difference	178	/noct-	"wild type Ser subst., with Ala"
F19	Misc-difference	187	/noct-	"wild type Ser subst., with Pro"
F20	Misc-difference	195	/noct-	"wild type Cys subst., with Asn"
F21	Misc-difference	202	/noct-	"wild type Gln subst., with Ala"
F22	Misc-difference	205	/noct-	"wild type Arg subst., with Lys"
F23	Misc-difference	208	/noct-	"wild type Gln subst., with Glu"
F24	Misc-difference	237	/noct-	"wild type His subst., with Asp"
F25	Misc-difference	243	/noct-	"wild type Asp subst., with Gly"
F26	Misc-difference	244	/noct-	"wild type Gln subst., with Lys"
F27	Misc-difference	249	/noct-	"wild type Glu subst., with Thr"
F28	Misc-difference	250	/noct-	"wild type Ser subst., with Pro"
F29	Misc-difference	261	/noct-	"wild type His subst., with Gln"
F30	Misc-difference	263	/noct-	"wild type Ala subst., with Ile"
F31	Misc-difference	266	/noct-	"wild type Ser subst., with Lys"
F32	Misc-difference	267	/noct-	"wild type Ala subst., with Pro"
F33	Misc-difference	275	/noct-	"wild type Lys subst., with Glu"
F34	Misc-difference	282	/noct-	"wild type Thr subst., with Lys"

FT Misc-difference 287 /note- "wild type Ala substd. with pro"
 FT Misc-difference 400 /note- "wild type Gly substd. with Asn"
 FT Misc-difference 401 /note- "wild type Leu substd. with Phe"
 FT Misc-difference 329 /note- "wild type Glu substd. with Asp"
 FT Misc-difference 344 /note- "wild type Ala substd. with Val"
 FT Misc-difference 363 /note- "wild type Asn substd. with Ser"
 FT Misc-difference 371 /note- "wild type Thr substd. with Ser"
 FT Misc-difference 395 /note- "wild type Leu substd. with Pro"
 FT Misc-difference 395 /note- "wild type Ser substd. with Gly"
 FT Misc-difference 401 /note- "wild type Leu substd. with Pro"
 FT Misc-difference 403 /note- "C-terminal addition to the wild type sequence"
 FT W952639-A1.
 EN 05-OCT-1995.
 XX 24-MAR-1995: 95WO-JP00541.
 XX 04-SEP-1994: 94JP-0216333.
 PR 25-MAR-1994: 94JP-0056271.
 XX (MAKU) MAKUKIN SHOUU KK.
 FT Matsu I, Ohba Y, Tsukada Y:
 OR WP1: 1995-351320/45.
 XX Recombinant acylglucosamine-2-epimerase with renin-binding activity
 FT - useful in enzymic production of N-acetylmannosamine and
 FT N-acetylneuraminic acid
 PS Claim 13: Pages 59-61; 74pp: Japanese.
 XX R79929 is a porcine acylglucosamine-2-epimerase (AZP) mutant,
 CC with renin-binding activity. AZP can be used for the enzymatic
 CC prodn. of N-acetylmannosamine and N-acetylneuraminic acid.
 XX Sequence 417 AA:
 SN
 Query Match 30.0%; Score 645; DH 16; Length 417;
 *Best Local Similarity 34.7%; Pred. No. 1.3e-59;
 Matches 138; Conservative 75; Mismatches 163; Indels 22; Gaps 7;
 5 BRQELAQGVYVALHDD---VLPFWKYSIDRQGGYFCIDRRKGOVPIIDKFLIMQNGOV 61
 3 KRELGLGKCKVGGQIDIVVAFWKSHDQGLGQIFTELFGQGVYDGLKYWLIQRTGY 62
 62 WQFAVYNNLE--PKQWLEIRHSGADFLARHGR--DQDKNYFALQDEKPLRQPNVF 117
 64 WYGLTYLTFITFIHQGLIDAAKAGGCFILYATVAPGPKKCFVLTIDGIPKVGRTIF 122
 118 SUGPAAHATSGVATASGADFAKALIAQAVNNVR-KOHNPKGQYKSYGSTRPLKSLAVP 176
 123 SECTYMMNMLTWRALQEVRYQLQAVEMMDQIVHWYQDASQIQRPIQAPAAPMAVP 182
 177 MLIANILFEMNLTPTVEVLAQTVREV-----MDPILDPHGLGMRKPAVTPIGREV 229
 184 IMLIHLVGGI-----QADQGLQAGYACIQDWCAITLIQHWYFDQAGVLENVSEGGKEI 246
 240 DSFSELLNPGSHIEAMWEMMDIAQRSSDRQLQEGAI-AVVLNTLEYAVDEDFGIFYEI 288

DB 237 pncIarqapnphlleaawllllhrlkqdpelrahvldklllpilssawp-dhaqllyh 296
 QY 289 DRQGHPPQGLFEMQKLMWVHLETLVALAKQDAIDQEKWQWREHVDYAMSHFAPREV 348
 DB 297 daahncpdlqewamkLwphseamidlmsysdsqplvllrlyvaeylrlqrdpeya 456
 QY 349 EMEGYLNRKGEVLLNKGSRKWKSCFFVPRALWCAETL 386
 DB 357 ewtgyistegkvalstlkypfkqctivprclamecmml 394
 RESULT 5
 ID Y39296 standard; protein; 532 AA.
 XX Y39296:
 AC Y39296:
 XX 02-DEC-1999 (first entry)
 DT
 XX Alkaline Bacillus alpha-amylase.
 DE
 XX Alpha-amylase; detergent; laundry; hard surface cleaning; sweetener;
 KW ethanol; starch conversion; liquefaction; saccharification; pulp;
 KW paper; cardboard; lignocellulosic material; textile desizing; alkaline;
 XX beer making; mashing process.
 OS
 XX Bacillus sp.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..31
 FT /label- Signal_peptide
 FT Protein 32..532
 FT /label- Alpha-amylase
 FT /note- "The mature protein is specifically claimed"
 XX W09942567-A1.
 XX 26-AUG-1999.
 PD 15-FEB-1999: 99WO-0K00066.
 XX 18-FEB-1998: 98DK-0000228.
 PR (NOVO) NOVO-NORDISK AS.
 PA Outtrup H, Nielsen BR, Hedegaard L, Andersen JJ:
 PI WP1: 1999-550763/46.
 DR N-PSDB: 206797.
 XX Novel alpha-amylase enzymes derived from a strain of Bacillus
 PT Claim 1: Page 32-34; 39pp: English.
 XX
 CC This is a Bacillus species alpha-amylase amino acid sequence. The
 CC alpha-amylase is derived from a strain of Bacillus. The preferred strain
 CC is Bacillus sp. N01M8 40916. The mature protein sequence and the
 CC polynucleotide coding for the mature protein sequence are specifically
 CC claimed. The protein has a molecular weight of 55kD, as determined by SDS
 CC page, and an iso-electric point of about 5. The amylase has optimum
 CC activity at about 55 degrees C and pH 9.5, and can be used in the
 CC production of improved detergents, e.g. laundry and hard surface cleaning
 CC detergent compositions. It is also useful in the production of sweeteners
 CC and ethanol from starch and also in the conventional starch-converting
 CC processes such as liquefaction and saccharification. Alpha-amylase may
 CC also be used in the production of lignocellulosic materials, such as
 CC pulp, paper and cardboard, from starch reinforced waste paper and
 CC cardboard. The new alpha amylase may also be used in modifying starch in
 CC the presence of alkaline fillers such as calcium carbonate, in the
 CC textile industry for textile desizing, and in beer-making during the
 CC mashing process.


```

UY 64 FAVYNNEL-----EPKQVMEIARHG-ADPEARHGRDQ-----GNWVF-ALDQE 106
DB 173 ----IqrIyKrkKwkwvdsuqdyIImyadvdhndpvevneIrrwqewyInIld 229
UY 107 G-----KPIKQRYN-----VPSDFAMAEsuytLAsIvOeAKRAlALuANNVL 150
DB 240 qftIdavKhkIystIdwtIthvmtatKontvavctwknIdq-----aleuyInkt 280
UY 151 RROI-----NPKIvOYEKSYPGTIEPLKSLAVPMILANLEMEILLEPTTV 195
DB 281 dwkIsvidrphIhnyIAsnsqny-----dmakIInatv 316
UY 196 PEVIACTVREVMIDIDPEIGIMKEAVIPIGEVVDSFENALNPQCIEM---WPMMD1 252
DB 417 qK-----Ipr-----mbavt-----Ivdhdst-----pqsIestIvqwt----- 346
UY 253 AKRSRIRIOLOAOAVVNLNLEVMDEEFSGTIEFLRQGHPRQOUEWQKLMWILETL 312
DB 447 -----KPIyayIItI-rrcqIypsvTygydyqIphIsvpmkKIdp-----Ileat 391
UY 413 VALAKSHIGATQGEKQWQWFERVHDY-----AMSH-----FADPEVGEWF 351
DB 492 qutayayIq-----IdyIdhmlIqKreantIhpsqIalImstqpaqekw 437
UY 452 GYL---NRKQVILNIKKOK 368
DB 438 myzqpkKaqpwIdIqK 456

```

RESULT 11

ID W26684 standard: protein: 999 AA.

XX W26684:

DB 27-MAR-1998 (first entry)

XX

DE Bovine lysosomal alpha-mannosidase (LAMAN) R221H mutant.

XX

KM LAMAN: lysosomal alpha-mannosidase; alpha-mannosidosis; cattle;

KW diagnosis: screening; genetic test.

XX

OS Bos taurus.

XX

EH Key Location/Qualifiers

ET Modified site 134

ET Modified site /note- "N-glycosylated"

ET Modified site 499

ET Modified site /note- "N-glycosylated"

ET Modified site 499

ET Modified site /note- "N-glycosylated"

ET Modified site 644

ET Modified site /note- "N-glycosylated"

ET Modified site 640

ET Modified site /note- "N-glycosylated"

ET Modified site 681

ET Modified site /note- "N-glycosylated"

ET Modified site 755

ET Modified site /note- "N-glycosylated"

ET Modified site 919

ET Modified site /note- "N-glycosylated"

XX

EN W0926369-A1.

XX

DB 24-JUL-1997.

XX

XX 15-JAN-1997: 97WD-GH00109.

XX

XX 15-JAN-1996: 96ND-0000163.

XX

PA (HEB67) HEB67 T.

PA (HEA7) HEA7 T.

PA (NUS7) NUSSEN G.

PA (COL7) COLLERSKD OK.

PA (DZIE7) DZIEGLEWSKA H.

XX

PI Berg T, Nilsen G, Tollstrand OK;

XX

DB WP1: 1997-08-05/2/35.

DB N-PSDB: 191095.

XX

PT Diagnosis and screening for bovine alpha mannosidosis - by detecting

PT mutation(s) in alpha-mannosidase gene, also nucleic acid encoding

PT the enzyme and derived oligonucleotide primers

XX

PS Example 2: Page 7: 85pp: English.

XX

CC This glycosylated polypeptide comprises a mutant bovine lysosomal

CC alpha-mannosidase (LAMN) that causes bovine alpha-mannosidosis

CC (barn) in Galleyway. The R221H mutation results from a point mutation

CC of the LAMN coding sequence (see 191094). A claimed method for

CC diagnosis or screening for barn comprises detecting the presence or

CC absence of barn causing mutations in the LAMN gene; another point

CC mutation causing barn in Angus and related breeds of cattle has also

CC been identified. PCR primers (see 191098-99) are provided for use

CC in a claimed method for detecting these mutations. Compared with

CC known enzymatic methods, smaller samples are needed. DNA is stable

CC during transport and storage, and the test is more reliable and

CC quicker.

CC (NB: the sequence for the R221H mutant alpha mannosidase was

CC produced by adaptation of the native sequence (see W26682) provided

CC in the specification).

XX

SV Sequence 999 AA:

Query Match 41% Score 88: DB 1st: Length 999:

Best Local Similarity 16.4%: Pred. No. 4.2:

Matches 63: Conservative 44: Mismatches 132: Indels 144: Gaps 14:

```

UY 108 KPIKQRYNVPSDFAMAEsuytLAsIvOeAKRAlALuANNVILRHQHPKQD----- 159
DB 26 raltprIprIsstIvIlaapraaayktCPKvKpmIhvbIphIdhkwIKIvdy 85
UY 160 ----YKSYV-GIPIKSIAYVMILANL-----LMEWILPPIIVIEVIACTV 204
DB 86 fytIymIqIqayqyIdsvIsIlauptrItyvclatIstswqIfoatqKivctIv 145
UY 205 EVMTOFLDPEIGIMKEAVIPIGEVVDSFENALNPQCIEM---WPMMD1 252
DB 146 qutIcIafanqwmvdcartIhyatIdqntIrrIfoctIqsdprIvayhdprIdhstq 205
UY 242 -----GIEAWPFMDI AKRSRIRIOLOAOAVVNLNLEVMDEEFSGTIEFLRQGH 284
DB 206 asIlaImatIdqIIt-----qIdyqKkKvKkIImeywIasIsIkpptadIIsv 257
UY 285 FVPIIDRQHIFVQIEMDQ----- 302
DB 258 ----IpmIyppqyIcwmIcadvkwvdcItrspynakelvrtIktatqKtytkIdv 314
UY 303 -----KLMWILETLVALAKSHQ-ATG-----QEKQWQWFERVHDYAMS- 440
DB 315 mtmgstqyenaIwtIKnIdkIqIvnaqgratIvovIystasyIwtInkalswsv 474
UY 341 -----HEADPEVGEWQYINR 367
DB 375 KKadIIPyadqpmIwtqYIsst 397

```

RESULT 12

ID W26683 standard: protein: 999 AA.

XX W26683:

XX

DB 27-MAR-1998 (first entry)

CC A rat cDNA clone was isolated from a spleen library in lambda-gt10
 CC on the basis of homology to the human integrin alpha-d subunit
 CC gene. RACE and PCR amplification were used to obtain the complete
 CC rat alpha-d sequence given in 991715. Rat alpha-d 1 domain/human
 CC 1984 fusions were used to raise MAbs.

XX Sequence 1161 AA:

Query Match 4.1k; Score 87.5; DB 16; Length 1161;

Best Local Similarity 20.2k; Pred. No. 4.5; Mismatches 123; Indels 141; Gaps 19;

Matches 79; Conservative 49; Mismatches 123; Indels 141; Gaps 19;

101 FALDQKRLKPYNYVESDFAMAFSQAASQACAKAIAIAVNNVLRQHNRKGV 160
 155 LLDASQSLQDIAQMKD-----LKAIMEIASLISIMYSNLI----- 198

161 EKSPGTGRLKSLAVPMILANLLEMMWLP-----TVEVLE-----AQVREV 206
 199-KHITTEKRLDPSLVQGVIGLTITATGRTMGGTFSKQSGTSKAKIILV 257

207 MTD-----FLDP-----ELGMEAVPTGEVDSFEGRLN-----FGHTEA 245
 258 IDGQKYPDLYSDVIPADKAGITRYAIGVQDATGPTAKKOLNIGSAPGDHVKV 317

246 MFMMDIAQRSQDQLOEQAIAV-----VINTL-EY 275
 318 QNL-----ADLSIQGLQPKLIDGLQSSSSSIQHEMSQGLSSALISDPVQAGVSL 374

276 AMDEFEGGILFYLDRQCHP-----QQLFMDQKIMWVHLFTIVALAKG----- 418
 475 SWS-----QAALY-----PPTLPRLIMNSQVMDSDYLYGSTAVAFKGVDSILIG 425

319 -HQAIAQ-----EKQWQFERVHDYAWSHFADPEYGEWGY-----LNKREVELLNK 465
 426 PRHGHTGKVVITFGQATHWPK-----SECVRTGISYQASISCVSDVDRDQSDVIL 479

466 G-----GKMKGCFHP--FALMLCAETL 486
 480 QAPHYEQRLDQGVSVLPDPAVGRWPEATL 511

RESULT 14
 W24062
 ID W24062 standard; Protein: 1161 AA.
 XX
 AC W24062;
 XX
 101- 24 FEB-1998 (first entry)
 XX
 DE Rat beta 2 integrin alpha d subunit.
 XX
 KW beta 2 integrin alpha d subunit; rat; cell migration;
 KW cell adhesion; phagocytosis; diabetes; atherosclerosis;
 KW multiple sclerosis; asthma; psoriasis; lung inflammation;
 KW acute respiratory distress syndrome; rheumatoid arthritis;
 KW hyalidoma; monoclonal antibody.
 XX
 KS Rattus sp.
 XX
 FH Key Location/Qualifiers
 F1 Misc-difference 474
 FT Misc-difference 1128 /note- "encoded by ACY"
 FT /note- "encoded by TYC"
 XX
 FN W09731099-A1.
 XX
 F0 28-AUG-1997.
 XX
 PE 24-FEB-1997; 97WD-0802713.
 XX
 PR 22-FEB-1996; 96US-0605672.

XX (100%) 100% C00P.

PI Gallatin MM. Van der Vliet M.

XX WPI: 1997-435154/40.

DE N-PSDB: T79257.

PT Hyalidoma 199M and antibody secreted by it. Specific for non-rat
 PT beta2 integrin subunit, useful to detect subunit in cells and
 PT modulate its activity

PS Example 17: Page 170-175; 222pp; English.

CC This polypeptide comprises the rat homologue of a novel human beta
 CC 2 integrin subunit, designated alpha d (see W24062). Its sequence
 CC was deduced from a cDNA clone (see T79257) isolated from a spleen
 CC cDNA library. Alpha d is involved in cell migration, phagocytosis
 CC and cell-cell interaction. Recombinant alpha d polypeptides can be
 CC expressed in transformed host cells for use in assays for
 CC identifying antibodies or other compounds that modulate alpha d
 CC activity or which modulate the interaction between alpha d and a
 CC ligand, for treating or preventing diseases in which macrophages are
 CC implicated. A monoclonal antibody that is specific for rat alpha d
 CC subunit, and which is secreted by hybridoma 199M (ATCC HB 12058),
 CC is claimed. It may be used to modulate alpha d activity.
 CC interfere with human alpha d, and treat or prevent diseases such as
 CC type 1 diabetes, atherosclerosis, multiple sclerosis, asthma,
 CC psoriasis, lung inflammation, acute respiratory distress syndrome
 CC and rheumatoid arthritis.

XX Sequence 1161 AA:

Query Match 4.1k; Score 87.5; DB 16; Length 1161;

Best Local Similarity 20.2k; Pred. No. 4.5;

Matches 79; Conservative 49; Mismatches 123; Indels 141; Gaps 19;

101 FALDQKRLKPYNYVESDFAMAFSQAASQACAKAIAIAVNNVLRQHNRKGV 160
 155 LLDASQSLQDIAQMKD-----LKAIMEIASLISIMYSNLI----- 198

161 EKSPGTGRLKSLAVPMILANLLEMMWLP-----TVEVLE-----AQVREV 206
 199-KHITTEKRLDPSLVQGVIGLTITATGRTMGGTFSKQSGTSKAKIILV 257

207 MTD-----FLDP-----ELGMEAVPTGEVDSFEGRLN-----FGHTEA 245
 258 IDGQKYPDLYSDVIPADKAGITRYAIGVQDATGPTAKKOLNIGSAPGDHVKV 317

246 MFMMDIAQRSQDQLOEQAIAV-----VINTL-EY 275
 318 QNL-----ADLSIQGLQPKLIDGLQSSSSSIQHEMSQGLSSALISDPVQAGVSL 374

276 AMDEFEGGILFYLDRQCHP-----QQLFMDQKIMWVHLFTIVALAKG----- 418
 375 SWS-----QAALY-----PPTLPRLIMNSQVMDSDYLYGSTAVAFKGVDSILIG 425

319 -HQAIAQ-----EKQWQFERVHDYAWSHFADPEYGEWGY-----LNKREVELLNK 465
 426 PRHGHTGKVVITFGQATHWPK-----SECVRTGISYQASISCVSDVDRDQSDVIL 479

466 G-----GKMKGCFHP--FALMLCAETL 486
 480 QAPHYEQRLDQGVSVLPDPAVGRWPEATL 511

RESULT 15

W72824

ID W72824 standard; Protein: 1161 AA.

XX
 AC W72824;

GenCorp version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

10M protein - protein search, using SW model

Run on: January 1, 2001, 22:22:04 : Search time 59.08 seconds
(without alignments)
94,870 Million cell updates/sec

Database: US-09-645-321-1

Percent scores: 21.47
1 MAAHRRQDIAQYVVALHQB.....CPHVKALMFAELLQGLVPS 391

Sequencing: HCSUM62
Gapexp 10.0 , Gapext 0.5

Searched: 164575 seqs, 1676186 residues

Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued: Patents_AA:*
1: /cqn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cqn2_6/prodata/2/1aa/5H-COMB.pep:*
3: /cqn2_6/prodata/2/1aa/6-COMB.pep:*
4: /cqn2_6/prodata/2/1aa/PTMS-COMB.pep:*
5: /cqn2_6/prodata/2/1aa/backlist1.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	694	42.3	402	1	US-08-553-703A-4
2	694	42.3	402	2	US-09-006-021-4
3	691	42.2	402	1	US-08-553-703A-1
4	691	42.2	402	2	US-09-006-021-1
5	668	41.1	419	1	US-08-553-703A-3
6	668	41.1	419	2	US-09-006-021-3
7	645	40.0	417	1	US-08-553-703A-2
8	645	40.0	417	2	US-09-006-021-2
9	87.5	4.1	1161	1	US-08-485-618-55
10	87.5	4.1	1161	2	US-08-485-618-55
11	87.5	4.1	1161	2	US-08-605-672-55
12	87.5	4.1	1161	2	US-08-482-293A-55
13	87.5	4.1	1161	2	US-08-943-463-55
14	86	4.0	485	2	US-08-446-803-2
15	86	4.0	485	2	US-08-861-847-2
16	86	4.0	485	3	US-08-600-656-2
17	84.5	3.9	729	1	US-08-070-165F-6
18	84.5	3.9	729	2	US-08-885-418-6
19	84.5	3.9	731	1	US-08-070-165F-10
20	84.5	3.9	731	2	US-08-885-418-10
21	81	3.8	357	2	US-08-476-254-11
22	81	3.8	359	5	5474933-8
23	79.5	3.7	1155	1	US-08-286-889-46
24	79.5	3.7	1155	1	US-08-485-618-46
25	79.5	3.7	1155	2	US-08-362-652-46
26	79.5	3.7	1155	2	US-08-605-672-46
27	79.5	3.7	1155	2	US-08-482-293A-46
28	79.5	3.7	1155	2	US-08-943-463-46

29	79.5	4.7	1161	1	US-08-485-618-53	Sequence 53, Appl
30	79.5	4.7	1161	1	US-08-462-652-53	Sequence 53, Appl
31	79.5	4.7	1161	2	US-08-605-672-53	Sequence 53, Appl
32	79.5	4.7	1161	2	US-08-482-293A-53	Sequence 53, Appl
33	79.5	4.7	1161	2	US-08-943-463-53	Sequence 53, Appl
34	78.5	3.7	509	2	US-08-557-122A-29	Sequence 29, Appl
35	78.5	3.7	1151	1	US-08-286-889-37	Sequence 37, Appl
36	78.5	3.7	1151	1	US-08-485-618-37	Sequence 37, Appl
37	78.5	4.7	1151	1	US-08-462-652-37	Sequence 37, Appl
38	78.5	4.7	1151	2	US-08-605-672-37	Sequence 37, Appl
39	78.5	3.7	1151	2	US-08-482-293A-37	Sequence 37, Appl
40	78.5	3.7	1151	2	US-08-943-463-37	Sequence 37, Appl
41	78	3.6	616	1	US-08-385-370-2	Sequence 2, Appl
42	78	3.6	616	1	US-08-385-370-4	Sequence 4, Appl
43	77.5	3.6	129	3	US-08-722-126A-10	Sequence 10, Appl
44	77.5	3.6	129	4	PCT-US95-04258-10	Sequence 10, Appl
45	76.5	3.6	773	1	US-08-019-870-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-553-703A-4
Sequence 4, Application US/08553703A
Patent No. 5795767
GENERAL INFORMATION:
APPLICANT: MARD, ISATIMI
APPLICANT: MITA, YASUHIRO
TITLE OF INVENTION: PHOSPHORASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knibbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPILED BY: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTA Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553, 703A
FILING DATE: 30-Nov-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/KEY NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-553-703A-4

Query Match 32.38 Score 694 Db 1 Length 402
Host Local Similarity 34.88 Prod. No. 73A-65
Matches 143 Conservative 77 Mismatches 153 Indels 46 Gaps 6

```

01 5 RIGELADVOYVQALNID VLEFWEKYSLABOGGAYPTVTLIRKQOVETPIREIWMLOKOV 61
02 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
03 8 KERELECAMKREWVLELEIRVAVFMLEHSHDEHJ24FETDEREKREYVUUKYVMIJCKOV 62
04 62 WOFVAVYVLE FRIOWLEFARICAOFLARBCF IOWNMVYFALDGBCEKELIGUPNVF 117
05 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
06 64 WOFVAVYVLEFRIOWLEFARICAOFLARBCF IOWNMVYFALDGBCEKELIGUPNVF 122
07 118 SOWVAVVAVSVALASVAGVAKALVAVNVVR KOHNRKQYREKSVYGLRPLSTAVP 174
08 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
09 124 SBTYVIMANMELMWTAVARVOSVAVEMRQULVIMVGRHSGLARPOLVAVSVASVAVP 182
10 177 MTLANMTEM EMLFTTYEVVAVVREVMVMTLAPELGLMREA 227
11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
12 183 RMUJLVVGLVGLVDELEFARVAVOJ2HW GARRLIGHVGRQA 235
13 222 VFDQEVVSEBRLINRCHITLFAWMTGMDVAVORSJHOLGEOAL AVULNELFVAMDE 280
14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
15 229 VGRHDEHLSVGLRPLNIGDALEFMLELIRHSKSGVAKLRAVYITFTTLFTFSWAD 286
16 281 FQALFVPLDQCHDQVGLQVLEWQKMWVHLELVAVAKCHQVAVQKRWOMFVAVVAMS 340
17 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
18 289 VGLTYVGLVAVDTTQLEWAMKLMWGRVGMALFELMYSRSHALRLLFYVAVETFR 348
19 341 HIAFVEVDEWVYVYINRGEVYVILNRKQKRWKQVTVYVAVMVGAVETLQVVS 401
20 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
21 349 GPHQVYVBMVYVYINRGEVYVILNRKQKRWKQVTVYVAVMVGAVETLQVVS 409
22
23 RESULT 2
24 US 09 006 021 4
25 Stepover 4, Application US/09006021
26 1 Patent No. 5994105
27 2 GENERAL INFORMATION:
28 3 APPLICANT: MARCO, ISATIMI
29 4 APPLICANT: GILIA, YASHIRO
30 5 APPLICANT: ISOKADA, YOSI
31 6 TITLE OF INVENTION: REIMERASE
32 7 NUMBER OF SEQUENCES: 9
33 8 CORRESPONDENT ADDRESS:
34 9 ADDRESSEE: Kudo, Matsuo, Osada & Imai
35 10 STREET: 620 Nopont Road, 16th Floor
36 11 CITY: Nopont Beach
37 12 STATE: CA
38 13 COUNTRY: U.S.A.
39 14 ZIP: 92660
40 15 COMPUTER READABLE FORM:
41 16 MEDIUM TYPE: Diskette
42 17 COMPUTER: IBM compatible
43 18 OPERATING SYSTEM: DOS
44 19 SOFTWARE: FastSeq Version 1.5
45 20 CURRENT APPLICATION DATA:
46 21 APPLICATION NUMBER: US/09/006, 021
47 22 FILING DATE:
48 23 CLASSIFICATION:
49 24 PRIOR APPLICATION DATA:
50 25 APPLICATION NUMBER: 08/554,704
51 26 FILING DATE: 30 Nov 1995
52 27 ATTORNEY/AGENT INFORMATION:
53 28 NAME: Altman, Daniel E.
54 29 REGISTRATION NUMBER: 44,115
55 30 REFERENCE/DOCID NUMBER:
56 31 TELECOMMUNICATION INFORMATION:
57 32 TELEPHONE: 714 760 0404
58 33 TELEFAX: 714 760 9502
59 34 FAX:
60 35 INFORMATION FOR SEQ ID NO: 4:
61 36 SEQUENCE CHARACTERISTICS:
62 37 LENGTH: 402 amino acids
63 38 TYPE: amino acid
64 39 STANDARDINSET: single
65 40 TOPLOGY: linear

```

```

1      MOLECULE TYPE: peptide
2      FRAGMENT TYPE: N-terminal
3      OS: 09-006-021-1
4
5      Query Match: 32-38; Score: 654; Ids: 2; Length: 402;
6      Host Local Similarity: 34.8%; Prod. No.: 7,40,65;
7      Method: 14; Conservation: 77; Mismatches: 153; Indels: 48; Gaps: 8;
8
9      1 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
10     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
11     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
12     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
13     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
14     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
15     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
16     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
17     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
18     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
19     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
20     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
21     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
22     1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	32/28	Score 693	EB 1	Length 402
Host local Similarity	85/08	Pred. NO.	1.5e-64	
Matches	144	Conservative	74	Mismatches 155
				Models 48
				Gaps 8

07	5	KROGELADYVVALIND	---VLFERKESILTRUGASVITLIDKROKVOITIKELIMOROV	6
08	4	KRETELAKMERKVOGJEDPRVAMFLESHIDEIRHJEDFTCLASROKRYVADOLKYVALDROGV	6.2	
07	6.2	WOFAPVFNKLE---	IKTOMELIARIGADPELAHDK--WGDONMVFALDROKPIKROPVNF	117
08	6.3	MMYCKLYKKELEFRIFERILLIDAKKAGSELLIKHARVAPPEKKAFELTRIDRPPKVSJISF	122	
07	118	SUGTAAAMFVSVALSNOEKKATVALOVANNVLR	KOHNRKQYKESVOTIFREPLKSLAP	176
08	123	SECFEIMAMNFIWVLAHAKYQSAVIMMOJIVHWVHDHDSIGLRIPIQAVASASMAVP	182	
07	127	MLIANIILEM---	EMILPTVEVLAOTVREVMIFILPELIGREK	221
08	184	MMMLLTIVGULDEELBELLAGKRYOLGJIM	--CARILLIUVIGROQA-----VLEN	228
07	2.22	VILPDIQVHSPGKLEINHGHCIFAMFPMIDIAOSGSHRQJOFQAM	-AVVINILPEAMDE	280
08	2.29	VSEMEHELISGCLGRHJONGHLEAMSWELLIKRISSTAKLEABVIDTELLIPERSMDAD		288
07	281	FUGITVILLDROGHIFPOULEMIOUKLWVILETTVALAKSHOATVQERK	WUMFERHIDYAS	340
08	289	HOGIIFYVQALANDITIGULEMARKIMWGHSHAMIAFIMQYSHSGPDAIIRQYVAVATFR		348
07	341	HHADPEYEMFVYENRGESEVILNKGKWKQCEIVAFALMIACTDOLPUS	391	
08	349	QFROPPEYEMFVYENRGEKVALITKROGPFKGTIVPR	LAMCEEMLSALLS	399

RES-11-4
US-09-006-021-1
Sequence 1, Application US/09006021

1 APPLICATION NUMBER: 08/554,703
 2 FILING DATE: 30-MAY-1995
 3 ATTORNEY/AGENT INFORMATION:
 4 NAME: Altman, Daniel E.
 5 REGISTRATION NUMBER: 34,115
 6 REFERENCE/ID NUMBER:
 7 TELECOMMUNICATION INFORMATION:
 8 TELEPHONE: 714-760-0404
 9 TELEFAX: 714-760-9502
 10 FILE:
 11 INFORMATION FOR SEQ ID NO: 1:
 12 SEQUENCE CHARACTERISTICS:
 13 LENGTH: 402 amino acids
 14 TYPE: amino acid
 15 STRANDEDNESS: single
 16 TOPOLOGY: linear
 17 MOLECULE TYPE: peptide
 18 FRAGMENT TYPE: N-terminal
 19 US-07-006-021-1

Query Match	42.28	Score	691	106.2	Length	402	
Best Local Similarity	35.08	Pred. No.	1.55	64			
Matches	144	Conservative	74	Mismatches	155	Indels	40
						Gaps	8

[illegible]

```

1 RESULT 5
2 US-08-553-703A-3
3 Sequence 3, Application US/08553703A
4 Patent No. 5795767
5
6 GENERAL INFORMATION:
7
8 APPLICANT: MAKU, ISAFUMI
9
10 APPLICANT: OHEA, YASUHIRO
11
12 APPLICANT: TSUKADA, YOTI
13
14 TITLE OF INVENTION: ENHANCE
15
16 NUMBER OF SHOOTINGS: 4
17
18 CORRESPONDENCE ADDRESS:
19
20 ADDRESSEE: Knoble, Martens, [J]son & [J]son
21
22 STREET: 620 Newport Center Drive
23
24 City: Newport Beach
25
26 STATE: CA
27
28 COUNTRY: U.S.A.
29
30 ZIP: 92660
31
32 COMPUTER READABLE FORM:
33
34 MEDIUM TYPE: Diskette
35

```

```

1  COMPUTER: IBM Compat Ible
2  OPERATING SYSTEM: DOS
3  SOFTWARE: FastSeq Version 1.5
4  CURRENT APPLICATION DATA:
5  APPLICATION NUMBER: US/08/554,703A
6  FILING DATE: 30 Nov 1995
7  CLASSIFICATION: 45
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER:
10 FILING DATE:
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Attorno, Daniel E.
13 REGISTRATION NUMBER: 44,115
14 REFERENCE/PACKET NUMBER:
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 714 760 0404
17 TELEFAX: 714 760 9502
18 INDEX:
19 INFORMATION FOR SEQ ID NO: 4:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 419 amino acids
22 TYPE: amino acid
23 STRANDNESS: single
24 TOPOLOGY: linear
25 MOLECULE TYPE: peptide
26 FRAGMENT TYPE: N-terminal
27 US:09-645-321-1

```

```

Query Match: 41.18% Score 608; ID 1; Length 419;
Best Local Similarity: 45.68%; Pred. No. 4; Gap 62;
Matches: 145; Conserved: 71; Mismatches: 151; Indels: 40; Gaps: 9;

```

```

1  5  RQGLADQYVGLHD  VLPRWKYSIDRGZNYFTGLRRGVYDTIRKTLQNGV 61
2  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
3  4  KRELLGVKKQVQVGLSVLAFWMSHSDRGCFGLRRGVYDTIRKTLQNGV 62
4  6 2  WQFAVYNNLE  PRQVLELARIQALRGR  PRGRNNYFALDGRKTLQNGV 117
5  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
6  6 3  WMYVFLYFRFKEVVELDAKAGREFLSYAKVAPGRKAVFLYGRGVVGRITF 122
7  118  SPTFAAPSOYALASQAEKAKALQAVNNVLR  KOHNKQYKSYVCIIRKSLAV 176
8  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
9  124  SPTFYMMNMLKVTGEMHYGRVAVRMQDILHWVRELFAGVQSLTALERMV 182
10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
11 177  MLANLLEKMLPPTVEVLAQTVKVMIDELPRG  -----LMRE 220
12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
13 183  MLNLT  VQGLGRD  FEMTKY  AELDWM  AHRLLGHVGRDQVLE 227
14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
15 221  AVLPGRVSTSRDLNINSHGLEAMWPMMDIAGSRDQLQVAL  AVANTLEYAGD 279
16 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
17 228  NVSHKRLDPTLDQNLNCHLEAWFLDYALRGRDRIQRI  DKLTLLEFHSMD 287
18 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
19 280  EPGDLYEFLDGRGTPVGLQEMQKIMWVHLELVALAKHQALQGRKQWMPFRVHYAW 339
20 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
21 286  EPGDLYEFLDGRGTPVGLQEMQKIMWVHLELVALAKHQALQGRKQWMPFRVHYAW 347
22 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
23 340  SHPADEYGVMPVYNRGRVILNKKGRKKGTTHVPRALMWAPL 386
24 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
25 348  HGRDQVGVMPVYNGRQVALTRGRGRKTRVHPRALMWAPL 394

```

```

1  RESULT:
2  US:09-645-321-1
3  Sequence: 4; Application: US/0906021
4  Patent No.: 5901105
5  GENERAL INFORMATION:
6  APPLICANT: MARI, ISAORI
7  APPLICANT: OHIA, YASUHIRO
8  APPLICANT: ISIKAWA, YOSHIO
9  TITLE OF INVENTION: ESTERASE
10 NUMBER OF SEQUENCES: 9
11 CROSS-REFERENCE ADDRESS:

```

```

1  ADDRESS: Knudsen, Mattens, Olson & Best
2  STREET: 620 Newport Center Drive, Fifth Floor
3  CITY: Newport Beach
4  STATE: CA
5  COUNTRY: U.S.A.
6  ZIP: 92660
7  COMPUTER READABLE FORM:
8  MOBILE TYPE: Diskette
9  COMPUTER: IBM Compat Ible
10 OPERATING SYSTEM: DOS
11 SOFTWARE: FastSeq Version 1.5
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/09/006,021
14 FILING DATE:
15 CLASSIFICATION:
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: 08/554,703
18 FILING DATE: 30 Nov-1995
19 ATTORNEY/AGENT INFORMATION:
20 NAME: Attorno, Daniel E.
21 REGISTRATION NUMBER: 44,115
22 REFERENCE/PACKET NUMBER:
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE: 714 760 0404
25 TELEFAX: 714 760 9502
26 INDEX:
27 INFORMATION FOR SEQ ID NO: 4:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH: 419 amino acids
30 TYPE: amino acid
31 STRANDNESS: single
32 TOPOLOGY: linear
33 MOLECULE TYPE: peptide
34 FRAGMENT TYPE: N-terminal
35 US:09-645-321-1

```

```

Query Match: 31.18% Score 608; ID 2; Length 419;
Best Local Similarity: 35.68%; Pred. No. 4; Gap 62;
Matches: 145; Conserved: 71; Mismatches: 151; Indels: 40; Gaps: 9;

```

```

1  5  RQGLADQYVGLHD  VLPRWKYSIDRGZNYFTGLRRGVYDTIRKTLQNGV 61
2  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
3  4  KRELLGVKKQVQVGLSVLAFWMSHSDRGCFGLRRGVYDTIRKTLQNGV 62
4  6 2  WQFAVYNNLE  PRQVLELARIQALRGR  PRGRNNYFALDGRKTLQNGV 117
5  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
6  6 3  WMYVFLYFRFKEVVELDAKAGREFLSYAKVAPGRKAVFLYGRGVVGRITF 122
7  118  SPTFAAPSOYALASQAEKAKALQAVNNVLR  KOHNKQYKSYVCIIRKSLAV 176
8  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
9  124  SPTFYMMNMLKVTGEMHYGRVAVRMQDILHWVRELFAGVQSLTALERMV 182
10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
11 177  MLANLLEKMLPPTVEVLAQTVKVMIDELPRG  -----LMRE 220
12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
13 183  MLNLT  VQGLGRD  FEMTKY  AELDWM  AHRLLGHVGRDQVLE 227
14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
15 221  AVLPGRVSTSRDLNINSHGLEAMWPMMDIAGSRDQLQVAL  AVANTLEYAGD 279
16 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
17 228  NVSHKRLDPTLDQNLNCHLEAWFLDYALRGRDRIQRI  DKLTLLEFHSMD 287
18 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
19 280  EPGDLYEFLDGRGTPVGLQEMQKIMWVHLELVALAKHQALQGRKQWMPFRVHYAW 339
20 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
21 286  EPGDLYEFLDGRGTPVGLQEMQKIMWVHLELVALAKHQALQGRKQWMPFRVHYAW 347
22 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
23 340  SHPADEYGVMPVYNRGRVILNKKGRKKGTTHVPRALMWAPL 386
24 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
25 348  HGRDQVGVMPVYNGRQVALTRGRGRKTRVHPRALMWAPL 394

```

```

1  RESULT:
2  US:08-554-703A-2
3  Sequence: 2; Application: US/0854703A

```


Best Local Similarity 20.28; Pred. No. 2.8;
Matches 79; Conservative 49; Mismatches 123; Indels 141; Gaps 19;

```

QY 101 FAHQCKPLINQYVNNVSDCEPAAAFQSYALASCAQJAKAIALQAVNNVLRHOHNGQV 160
Db 155 FLIDGNSININQROFQOKKD-----FKALMGFEVASTLEISLMQYNIL----- 198
QY 161 EKSTPGTRPLKSLAVPMILANLTLEMWLLP-----TVEEVL-----AQVREV 206
Db 199 -KTHFETFEKNLIDPOSTVDPVLOJGLLIYTAGIKTWMELPLSHKSGSRKSAKILIV 257
QY 207 MTD---FLDP-----EILMRENTPTGEEFSDSEGLLN-----PGHTEA 245
Db 258 ITIDQKRYKDPLEYSQVLPADAKGIIIRYALGVSDAFQEPALTALKEINTGSAPPQDHFKV 317
QY 246 MAFPMMDIAORSGDPSOLQSOALAV-----VINIL-FY 275
Db 318 GNF---ALNRILQVLOLQEKITAIEGTQSRSSSTQJHEMSQDGHSSALTSKGPVLGAVGF 374
QY 276 AMDEEFGSIEVELDROGHP-----QJLEMDOKLMWHIETLVALAKG----- 318
Db 375 SMS---GQAFILY-----PPNTRPPLINMSQDANDMSQSLQSYSLAAAFKKQVHSLILGA 425
QY 319 --HQAATQ---EKQWMEERVIIDVAMSHFADQEVYEMQY-----LNRQEVILNKK 365
Db 426 PRHQHTGKVAVIFTEQARIMRER-----SEVRQTQIGSYFNASLQSDVDQDSXDLEVI 479
QY 365 G-----GKMKGCFHPV--RALMILCAETL 386
Db 480 GAPIIVYEDQTRAGQVSVFPPQVGRKQWCEATL 511

```

RESULT 11
US-08-605-672-55
? Sequence 55, Application US/08605672

```

: GENERAL INFORMATION:
:
: APPLICANT: Galatiu, W. Michael
:
: APPLICANT: Van der Vliet, Monica
:
: TITLE OF INVENTION: No. 581751 Sel Human 2 Integrin Alpha Subunit
:
: NUMBER OF SPOUNTERS: 103
:
: CORRESPONDENCE ADDRESS:
:

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

```

1 CLASSIFICATION: 530
2
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 08/173,497
5 FILING DATE: 21-DEC-1993
6
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: US 08/286,889
9 FILING DATE: 5-AUG-1994
10
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 08/362,652
13 FILING DATE: 21-DEC-1994
14
15 ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32688
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

```

: TELEX: 25-1856
: INFORMATION FOR SEQ ID NO: 55:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1161 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-605-672-55

```

Query Match	4.18	Score 87.5	DM 2	Length 1161
Best Local Similarity	20.28	Pred. No. 2.8		
Matches	79	Conservative	49	Mismatches 123
				Indels 141
				Gaps 19

[illegible]

RESULT 12
US-08-482-293A-55
; Sequence 55, Application US/08482293A
; Patent No. 5831029

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

STATE: Illinois
COUNTRY: United States
Zip: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:

1 FILING DATE: 08/28/89
 2 CLASSIFICATION: 530
 3 PRIOR APPLICATION DATA:
 4 APPLICATION NUMBER: US 08/173,497
 5 FILING DATE: 23-DEC-1993
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: US 08/286,889

APPLICANT: Rasmussen, Michael Dolberg
 APPLICANT: Van der Zee, Pia
 TITLE OF INVENTION: Alkaline Bacillus Amylase
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: No. 58245310 No. 58245310disk of No. 58245311th America
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,803
 FILING DATE: 01-June-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Harrington, James J.
 REGISTRATION NUMBER: 48,711
 REFERENCE/DOCKET NUMBER: 4157, 204-US
 TELEPHONE: (212) 867-0123
 TELEFAX: (212) 878-9655
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 485 amino acids
 TYPE: amino acid
 STANDARDS: single
 TOPOLOGY: Linear
 MOLECULE TYPE: peptide
 US-08-446-803-2

Query Match 4.08; Score 86; DB 2; Length 485;
 Best Local Similarity 20.08; Pred. No. 1.1;
 Matches 87; Conservative 46; Mismatches 127; Indels 174; Gaps 24;

07 4 HRRQLAQQYVQAIHQVLPFWKYSIDRGQGYFTCLRRKQVYDFDKETWLNQNVQW 63
 126 NRNDPISGV-----TFAMIKFDPGKNTY--SHKRWYHFGVMDQSKQ-FV 174
 05 64 FAVFYNRLEKRP-QWLEIARHG-ADFLARIGRQD-----GNWYF-ALDQES-- 107
 175 NRIYKFGDKKAWIWDVNSGNDYIMYADVMDHFEVNEIRKRWCHWYNTNLTNDGR 234
 07 108 ---KPIKQRYN-----VPSDFEAMAFSQAALASGAQFAKALAIQAYNNVLRQ 153
 10 245 IDAVNHITKYSTKRWLTIVRNATGKEMFAVAEFKKNLG-----ALENTLNKTNN 285
 07 154 H-----NPKQYKSYDQIKPKLSIAVPMILANIEMEMILPPIVHEV 198
 10 286 HSYVDVPLHYNNYNASNGNY-----DMAKLNGIVQK- 320
 07 199 LAQTVREVMTPFLPELIGLREAVTPTGEVDSFEGKLLNFGIGLEAM--WFMMDIAQK 255
 10 421 -----HP-----MHAVT-----FVDMHDSQ-----FGESLSFVQFWF----- 448
 07 256 SDQRLODQALAVVINTLEVAWDEFGQTFYFLDRQCHPQULMDQKLMWVHELEVAL 315
 10 449 -----KFLAVALL-TEGGYPSVFYGYGIGTITNSVPARKAKIDP-----LEAVRNF 396
 07 416 AKGHQATQGEKQWQWFERVHY-----AMSH-----FAQVYCHWQDYL 454
 10 497 AYGTD-----HNYFHHNLTQWLRQENLTIPNSQIALIRNSAQDQKRWYV 442
 07 455 ---NRGEVLLNKIKK 668
 10 443 QGNKAGVWHDITQNK 458

RESULT 15
 US-08-861-837-2
 Sequence 2, Application US/0881837
 Patent No. 5856164
 GENERAL INFORMATION:
 APPLICANT: Ottarp, Helge
 APPLICANT: Hisgard-Plantzow, Henrik
 APPLICANT: Osteraard, Peter Rabek
 APPLICANT: Rasmussen, Michael Dolberg
 APPLICANT: Van der Zee, Pia
 TITLE OF INVENTION: Alkaline Bacillus Amylase
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: No. 58561640 No. 58561640disk of No. 58561641th America
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/861,837
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/446,803
 FILING DATE: 01-June-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Harrington, James J.
 REGISTRATION NUMBER: 48,711
 REFERENCE/DOCKET NUMBER: 4157, 204-US
 TELEPHONE: (212) 867-0123
 TELEFAX: (212) 878-9655
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 485 amino acids
 TYPE: amino acid
 STANDARDS: single
 TOPOLOGY: Linear
 MOLECULE TYPE: peptide
 US-08-861-837-2

Query Match 4.08; Score 86; DB 2; Length 485;
 Best Local Similarity 20.08; Pred. No. 1.1;
 Matches 87; Conservative 48; Mismatches 127; Indels 174; Gaps 24;

07 4 HRRQLAQQYVQAIHQVLPFWKYSIDRGQGYFTCLRRKQVYDFDKETWLNQNVQW 63
 126 NRNDPISGV-----TFAMIKFDPGKNTY--SHKRWYHFGVMDQSKQ-FV 174
 05 64 FAVFYNRLEKRP-QWLEIARHG-ADFLARIGRQD-----GNWYF-ALDQES-- 107
 175 NRIYKFGDKKAWIWDVNSGNDYIMYADVMDHFEVNEIRKRWCHWYNTNLTNDGR 234
 07 108 ---KPIKQRYN-----VPSDFEAMAFSQAALASGAQFAKALAIQAYNNVLRQ 153
 10 245 IDAVNHITKYSTKRWLTIVRNATGKEMFAVAEFKKNLG-----ALENTLNKTNN 285
 07 154 H-----NPKQYKSYDQIKPKLSIAVPMILANIEMEMILPPIVHEV 198
 10 286 HSYVDVPLHYNNYNASNGNY-----DMAKLNGIVQK- 320
 07 199 LAQTVREVMTPFLPELIGLREAVTPTGEVDSFEGKLLNFGIGLEAM--WFMMDIAQK 255
 10 421 -----HP-----MHAVT-----FVDMHDSQ-----FGESLSFVQFWF----- 448

57	256	SCHOLIOGATIAVAVINILILEYAMDEPENTILYFILLKCHHIFQULFWOKLMMWHLETVL	315
146	349	KILAVALLILITRECHTUSVFOGDTGILITHSVFAMRAKLTIO	11EARNQF
57	416	AKCHGATIGUEKOWMEFVAVY	AMSH
146	457	AVGILY	HEFYHHNLTOWRENTTTHNSLATIMSHQWREKMY
57	455	NEPEVILINIKJOK	668
146	445	QGNKRDYWHHILITOK	458

Search completed: January 1, 2001, 22:52:51
Job time: 1847 Sec

THE UNIVERSITY OF CHICAGO

WISCONSIN
JAN 11 1967

Eukaryotes; Metazoa; Chordata; Crustacea; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Rasbora; Danio.

1 (pages 1 to 542)

Clark, M., Johnson, S., Lee, C., Lebrach, B., Lee, R., Li, F., Martin, M., Fiddy

Matches	69	Conservative	0	Mismatches	59	Indels	0	Gaps	0
UY	707	tgctaacccgaagacagcagcatgaaagccatgtgtatcatgatgacatctgcccaagcct	766						
Db	872	tggtctacctggcgctgttgacagcggttcgcgtgaaactaacagtaacggttcggaagca	931						
UY	767	cgcagcagtcctcaatttacaagagcaaacgcatctgcagatgagatgagacaccccgaaatag	826						
Db	942	aagcagcaaacctctcttcctggcagcagcatcctgcagctgaagccggaaacccgagcagatag	991						
UY	827	cctgagat	844						
Db	992	tcctgcat	999						
RE-SEQ-1	3								
ID	Q20383								
XX	Q20383	standard; DNA: 2214 BP.							
XX	Q20383:								
XX	16-APR-1992	(first entry)							
XX	ADH	complex structural gene (72,000 mol. wt. protein).							
DE	Alcohol dehydrogenase; acetic acid; fermentation; ss.								
XX	Acetobacter alfoacetigenes NH-24.								
OS	JF03266988-A.								
XX	27-MAY-1991.								
XX	26-MAR-1990:	90JP-0073440.							
XX	26-FEB-1990:	90JP-0042301.							
XX	26-MAR-1990:	90JP-0073440.							
XX	(NAKA-) NAKANO SHIEN KK.								
XX	WPI: 1992-019325/03.								
DR	P-PSDB: R20192.								
XX	Alcohol dehydrogenase complex structural gene - used in plasmid								
PT	and enhancing efficiency of acetic acid fermentation for								
PI	transformed acetic acid bacteria								
XX	Disclosure: Fig 3(1-3); 21pp; Japanese.								
PS	The gene encodes a protein of mol. wt. ca. 72,000. Acetobacter								
CC	transformed with the sequence can enhance the efficiency of acetic								
CC	acid fermentation. The ADH complex can be easily extracted from the								
CC	bacteria and purified and it can be used for the determination of an								
CC	alcohol.								
CC	See also Q20383-84, and -86-88.								
XX	Sequence 2214 BP; 452 A; 683 C; 662 G; 417 T; 0 other;								
NU	Query Match	2 9%;	Score 33.6;	DB 13;	Length 2214;				
	Best Local Similarity	53.98;	Pred. No. 0.9;						
	Matches 69;	Conservative 0;	Mismatches 59;	Indels 0;	Gaps 0;				
UY	707	tgctaacccgaagacagcagcatgaaagccatgtgtatcatgatgacatctgcccaagcct	766						
Db	872	tggtctacctggcgctgttgacagcggttcgcgtgaaactaacagtaacggttcggaagca	931						
UY	767	cgcagcagtcctcaatttacaagagcaaacgcatctgcagatgagatgagacaccccgaaatag	826						
Db	942	aagcagcaaacctctcttcctggcagcagcatcctgcagctgaagccggaaacccgagcagatag	991						
UY	827	cctgagat	844						
Db	992	tcctgcat	999						
UY	827	cctgagat	834						

DB	992	tctgagcat	999
RESULT	4		
ID	Q69001	standard; cDNA; 2751 BP.	
XX	Q69001:		
XX	17-APR-1995	(first entry)	
XX			
DE		Product of alternative splicing of human proto-oncogene in 4' mRNA.	
XX		Cadherin; proto-oncogene; cell adhesion molecule; ss.	
OS	Homo sapiens.		
FI	Key	Location/Qualifiers	
FT	CDS	115-2160	
TT		/tag-a	
XX	MW9414960-A.		
XX	07-JUL-1994.		
XX	23-DEC-1993;	G3MO-0S12588.	
XX	PR	29-DEC-1992;	GZUS-0998003.
PA	(DOHE-) DOHENT EYE INST.		
PI	Suzuki S;		
DR	WPt: I994-293849/36.		
P-PSDB:	R49144.		
PT	Polynucleotide sequences encoding new proto-oncogenes useful		
XX	for modulating natural binding and regulating activities.		
XX	Example; Page 99-102; 114pp; English.		
CC	Two full length human cDNAs corresp. to the partial cDNAs of		
CC	HDMN-42 and HDMN-43 (Q68981,Q68982) were isolated		
CC	from human fetal brain cDNA library. Several overlapping cDNA		
CC	clones were isolated with each probe including two cDNAs which		
CC	contained the putative entire coding sequences of two novel		
CC	proteins designated proto-oncogene-42 (pc42) and proto-oncogene		
CC	(pc43). The DNA and deduced AA sequences of pc42 are in		
CC	Q68997/R58906, while those of pc43 are in Q68998/R58907. Sequence		
CC	analysis of various overlapping proto-oncogene cDNA clones revealed		
CC	that some clones contd. unique sequences at the 3' end. The		
CC	sequences forming the boundaries of the 3' end regions are		
CC	consistent with the consensus sequence of mRNA splicing, suggesting		
CC	that these clones may corresp. to alternatively spliced mRNAs.		
CC	The DNA and AA sequences of one possible product of alternative		
CC	splicing of pc42 mRNA are given in Q68999/R58911. The DNA and AA		
CC	sequences of two possible products of alternative splicing of pc43		
CC	mRNA are respectively presented in Q69000/R58912 and Q69001/R49144.		
Sequence	2751 BP; 698 A; 715 C; 696 G; 642 T; 0 other;		
Query Match	2.9%; Score 33.6; DH 15; Length 2751;		
Best Local Similarity	45.7%; Pred No. 1;		
Matches 117;	Conservative 0; Mismatches 139; Indels 0; Gaps 0;		
OY	302 ttcattgaatcagaagaacaacccctactaacacctatacaattttttcgattgt	461	
DB			
DB	914 ttgcacaagtatttgtagggaccaccaagggaatatatttaaccttcgcacacacaa	973	
OY	362 tcgccgatcatgaccttatcataatagctttagccagtgaggcagcagaagctaaga	421	
DB			
DB	974 accgagcacgactctgcgaacattatccccctltagacactltagacccaqqaia	1033	

db 1752 AGAATATGAGGAGGAAATGATATACATGCGGGAGTGGGCTTCGAGATCAT 1701

RESULT 12

g14465/c

10 g14465 standard: hMA: 4496 bp.

XX

XX

XX g14465/

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 12 FEB 1999 (first entry)

XX EST clone BX50.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 12 FEB 1999 (first entry)

XX EST clone BX50.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

XX 20-MAY-1998 (first entry)
XX
XX
XX Human macrosc cDNA 2.
XX
XX Macro scavenger receptor; MacroscR; human; antibodies; treatment;
XX diagnosis; cardiovascular disease; macrophage; bacterial infection;
XX immune cell related host cell disorder; pancreatitis; organ failure;
XX endotoxaemia; agonist; antagonist; dysfunction; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1560
XX FT /ftatg - a
XX FT /product - hMacroSR
XX FT /note - "partial macro scavenger receptor sequence"
XX
XX EF088899 A2.
XX
XX 26-NOV-1997.
XX
XX 19-MAY-1997: 97EP-0303394.
XX
XX 04-FEB-1997: 970S-0794795.
XX 23-MAY-1996: 960S-0017699.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Adamon JE, Elshourbagy NA, Gross MS, Lysko PG;
XX WPI: 1998-001786/01.
XX F-PCR: W6748.
XX
XX Human macro scavenger receptor - used to develop products for
XX treating e.g. cardiovascular disorders, septic shock, bacterial
XX infections etc.
XX
XX Claim 4: Fig 4: 45pp: English.
XX
XX This partial cDNA sequence encodes a novel human Macro scavenger receptor
XX (MacroSR). This polypeptide and resulting antibodies can be used to
XX treat cardiovascular diseases, including atherosclerosis, hypertension
XX myocardial and cerebral infarction, angina, organ failure, stroke,
XX gangrene, and loss of function in the extremities. They can also be used
XX to treat or diagnose various macrophage and other immune cell related
XX host defense disorders, septic shock, pancreatitis, multiple organ
XX failure, endotoxaemia and infections caused by gram negative and
XX gram positive bacteria. Agonists and antagonists may also be used to
XX treat these diseases. The polynucleotide can be used to detect mutated
XX forms of hMacroSR which may be associated with a dysfunction and can
XX define a diagnosis of a disease (or susceptibility to a disease)
XX resulting from altered hMacroSR expression.
XX
XX Sequence 1560 BP; 489 A; 387 C; 533 G; 251 T; 0 other;
XX
XX
XX Query Match 2.78; Score 32.2; DB 19; Length 1560;
XX Best Local Similarity 49.18; Pred. No. 2.2;
XX Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
XX
XX 900 ggaacaaagctctggtggtacatttgaacacctgatttgcctagacgaagccacaa 959
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 450 ggtcttcaaatcacaagaagccatgaacatgacctgaatgacctgacctgacacaa 519
XX
XX 960 aatcaatgacaaagaataatgttgcacatgatttgaacagatccatgaatacctcgaag 1019
XX | | | | | | | | | | | | | | | | | | | | | |
XX 520 cctgctgaagaagagagagcgtatggaagagatgagacaagacccctcctggga 579
XX
XX 1020 tcatttcgcgcataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1072
XX | | | | | | | | | | | | | | | | | | | | | |
XX 580 ccccaaaagccatccagagatcaagaagagagagagagagagagagagagagagag 632

```

```

RESULT 15
V09774
ID V09774 standard; cDNA: 1703 BP.
XX
XX V09774;
XX
XX 20-MAY-1998 (first entry)
XX
XX Human macrosc cDNA 1.
XX
XX Macro scavenger receptor; MacroscR; human; antibodies; treatment;
XX diagnosis; cardiovascular disease; macrophage; bacterial infection;
XX immune cell related host cell disorder; pancreatitis; organ failure;
XX endotoxaemia; agonist; antagonist; dysfunction; ss.
XX
XX Homo sapiens.
XX
XX EF088899 A2.
XX
XX 26-NOV-1997.
XX
XX 19-MAY-1997: 97EP-0303394.
XX
XX 04-FEB-1997: 970S-0794795.
XX 23-MAY-1996: 960S-0017699.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Adamon JE, Elshourbagy NA, Gross MS, Lysko PG;
XX WPI: 1998-001786/01.
XX
XX Human macro scavenger receptor - used to develop products for
XX treating e.g. cardiovascular disorders, septic shock, bacterial
XX infections etc.
XX
XX Disclosure: Pages 31-32: 45pp: English.
XX
XX This cDNA sequence encodes a novel human Macro scavenger receptor
XX (MacroSR). This polypeptide and resulting antibodies can be used to
XX treat cardiovascular diseases, including atherosclerosis, hypertension
XX myocardial and cerebral infarction, angina, organ failure, stroke,
XX gangrene, and loss of function in the extremities. They can also be used
XX to treat or diagnose various macrophage and other immune cell related
XX host defense disorders, septic shock, pancreatitis, multiple organ
XX failure, endotoxaemia and infections caused by gram negative and
XX gram positive bacteria. Agonists and antagonists may also be used to
XX treat these diseases. The polynucleotide can be used to detect mutated
XX forms of hMacroSR which may be associated with a dysfunction and can
XX define a diagnosis of a disease (or susceptibility to a disease)
XX resulting from altered hMacroSR expression.
XX
XX Sequence 1703 BP; 437 A; 423 C; 567 G; 275 T; 1 other;
XX
XX
XX Query Match 2.78; Score 32.2; DB 19; Length 1703;
XX Best Local Similarity 49.18; Pred. No. 2.3;
XX Matches 89; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
XX
XX 900 ggaacaaagctctggtggtacatttgaacacctgatttgcctagacgaagccacaa 959
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX 435 ggtcttcaaatcacaagaagccatgaacatgacctgaatgacctgacctgacacaa 494
XX
XX 960 aatcaatgacaaagaataatgttgcacatgatttgaacagatccatgaatacctcgaag 1019
XX | | | | | | | | | | | | | | | | | | | | | |
XX 495 cctgctgaagaagagagagcgtatggaagagatgagacaagacccctcctggga 554
XX
XX 1020 tcatttcgcgcataatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1072
XX | | | | | | | | | | | | | | | | | | | | | |
XX 555 ccccaaaagccatccagagatcaagaagagagagagagagagagagagagagagag 607

```

Search completed: January 1, 2001, 22:23:45
Job Time: 12:56:58


```

1 TELEPHONE: 412/474-6400
2 TELEFAX: 412/474-0448
3 TELEX: 25 4856
4 INFORMATION FOR SEQ ID NO: 106:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 2751 base pairs
7 TYPE: nucleic acid
8 STRANDEDNESS: single
9 TOPOLOGY: linear
10 MOLECULE TYPE: cDNA
11 FEATURE:
12 NAME/KEY: CDS
13 LOCATION: 115..2160
14 US 08 453 702A 106

Query Match          2.98; Score 33.6; DB 3; Length 2751;
Host Local Similarity 45.78; Prod. No. 0.45;
Matches 117; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

1 302 11gccttgaatcaagaaggaacccctgcgtcaacccatatacgttttttcgaattgct 361
2 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
3 914 tttaaacgaattctgaataagagaccccaacggtgaattatttatctcttgaaacacaca 973
4 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
5 362 tgcgcgcataagcctttagtaaatatgccttaacgaattgagcgcgaagaactaaagcca 421
6 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
7 974 acctggcgcggggctggcgaacattatcttgaaacttgaaacggaatcttaacacaca 1033
8 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
9 422 ttgcctgcgaagcctacaaataacgtctcgaagcctcgaacaaatcccaaaagcgaatcgc 481
10 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
11 1034 aagctgcctgcggaccttcgaagacacccaacaccccaacgaatttatcaccgcaaacaca 1093
12 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
13 482 aagaagcctacgaatgactaagaacctcaaatccctcgaagcgaatgaattttagcca 541
14 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
15 1094 aagacggaatactcgaagagacacatttcgaagcttgcttggaagcttggaattggaatg 1153
16 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
17 542 acctcaccctggaat 557
18 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
19 1154 acacgcccgcagacat 1169

RESULT 11
PCT-US94 12588 106
Sequence 106, Application PC/TUS9412588
GENERAL INFORMATION:
APPLICANT: SUZUKI, SHUNTARO
TITLE OF INVENTION: Protocol for Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: Boston
STREET: 6400 Sears Tower, 234 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM pc compatible
OPERATING SYSTEM: pc/pos/ms dos
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12588
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/998,003
FILING DATE: 29 Dec 1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Gilda E.
REGISTRATION NUMBER: 45,402
REFERENCE/WORK NUMBER: 41811
TELECOMMUNICATION INFORMATION:

```

```

1 TELEPHONE: 412/474-6400
2 TELEFAX: 412/474-0448
3 TELEX: 25 4856
4 INFORMATION FOR SEQ ID NO: 106:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 2751 base pairs
7 TYPE: nucleic acid
8 STRANDEDNESS: single
9 TOPOLOGY: linear
10 MOLECULE TYPE: cDNA
11 FEATURE:
12 NAME/KEY: CDS
13 LOCATION: 115..2160
14 PCT US94 12588-106

Query Match          2.98; Score 33.6; DB 6; Length 2751;
Host Local Similarity 45.78; Prod. No. 0.45;
Matches 117; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

1 302 11gccttgaatcaagaaggaacccctgcgtcaacccatatacgttttttcgaattgct 361
2 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
3 914 tttaaacgaattctgaataagagaccccaacggtgaattatttatctcttgaaacacaca 973
4 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
5 362 tgcgcgcataagcctttagtaaatatgccttaacgaattgagcgcgaagaactaaagcca 421
6 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
7 974 acctggcgcggggctggcgaacattatcttgaaacttgaaacggaatcttaacacaca 1033
8 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
9 422 ttgcctgcgaagcctacaaataacgtctcgaagcctcgaacaaatcccaaaagcgaatcgc 481
10 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
11 1034 aagctgcctgcggaccttcgaagacacccaacaccccaacgaatttatcaccgcaaacaca 1093
12 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
13 482 aagaagcctacgaatgactaagaacctcaaatccctcgaagcgaatgaattttagcca 541
14 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
15 1094 aagacggaatactcgaagagacacatttcgaagcttgcttggaagcttggaattggaatg 1153
16 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111 111
17 542 acctcaccctggaat 557
18 1111 1111 1111 111 111 111 111 111 111 111 111 111 111 111 111
19 1154 acacgcccgcagacat 1169

RESULT 12
PCT-US95-08071-106
Sequence 106, Application PC/TUS9508071
GENERAL INFORMATION:
APPLICANT: SUZUKI, SHUNTARO
TITLE OF INVENTION: Protocol for Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: Boston
STREET: 6400 Sears Tower, 234 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM pc compatible
OPERATING SYSTEM: pc/pos/ms dos
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08071
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12588
FILING DATE: 23 Dec 1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/998,003
FILING DATE: 29 Dec 1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Gilda E.

```

1 RESULT
 2 US-07-998-003A-104
 3 Sequence 104, Application US/07998003A
 4 Patent No. 5643781
 5 GENERAL INFORMATION:
 6 APPLICANT: SUZUKI, Shintaro
 7 TITLE OF INVENTION: Protocol for in Materials and Methods
 8 NUMBER OF SEQUENCES: 107
 9 CORRESPONDENCE ADDRESS:
 10 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
 11 ADDRESSEE: Hicknell
 12 STREET: 20 South Clark Street
 13 CITY: Chicago
 14 STATE: Illinois
 15 COUNTRY: USA
 16 ZIP: 60604
 17 COMPUTER READABLE FORM:
 18 MEDIUM TYPE: Floppy disk
 19 COMPUTER: IBM PC compatible
 20 OPERATING SYSTEM: PC-DOS/MS-DOS
 21 SOFTWARE: Patent in Release #1.0, Version #1.25
 22 CURRENT APPLICATION DATA:
 23 APPLICATION NUMBER: US/07/998-003A
 24 FILING DATE:
 25 CLASSIFICATION: 435
 26 ATTORNEY/AGENT INFORMATION:
 27 NAME: No. 5643781 and Greta E.
 28 REGISTRATION NUMBER: 35,302
 29 REFERENCE/WORK NUMBER: 30903

RESULT 14
 US-08-453-2748-104
 Sequence 104, Application US/0845342748
 Patent No. 5663300
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: Photocurable Materials and Methods
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, C/Oolo, Gaston, Murray & Boutour
 STREET: 6300 Sears Tower, 213 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent to Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/453,2748
 FILING DATE: 30-MAY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5663300and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 32660
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using SW model

Run on: January 1, 2001, 22:51:35 ; Search time 76.08 Seconds

(without alignments)
104.191 Million cell updates/sec

Title: US-09-645-321-1

Percent score: 2147
1 MIAHRQELDAQYVQALHOD.....CFHVPALMLCAETQLPVS 391

Sequence: 1 MIAHRQELDAQYVQALHOD.....CFHVPALMLCAETQLPVS 391

Scoring table: HUSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87994 seqs, 31947931 residues
Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	694	32.3	402	1 RHP_PIG	P17560 sus scrofa
2	668	31.1	419	1 RHP_RAT	P51607 ratius norv
3	645	30.0	417	1 RHP_HUMAN	P51606 homo sapien
4	201	9.4	387	1 MANA_RHIME	P29954 rhizobium m
5	148	6.9	413	1 YHIS_ECOLI	P32140 escherichia
6	98.5	4.6	689	1 YVAL_HAOSU	P37512 bacillus su
7	97.5	4.5	1007	1 MA2H_FHICA	O66432 felis stive
8	96	4.5	657	1 HRY_PALVO	P80888 palinurus v
9	95	4.4	876	1 SYV_MYCTO	O63175 mycobacteri
10	94.5	4.4	1010	1 MA2B_HUMAN	O60754 homo sapien
11	93.5	4.4	2493	1 CYAA_USIMA	P49606 ustilago ma
12	93	4.3	657	1 HCYB_PANIN	P10787 pennitrus i
13	92.5	4.3	748	1 GUNC_HSIPL	P27033 pseudomonas
14	91	4.2	546	1 SYL_PPRFO	P46214 pyrococcus
15	91	4.2	657	1 HCYA_PANIN	P40254 panulirus i
16	90.5	4.2	582	1 HEMO_OPSA	P41090 opsanus lau
17	89.5	4.2	583	1 TRPA_HUMAN	O43280 homo sapien
18	88.5	4.1	698	1 GSGI_YEAST	P46944 saccharomyc
19	88	4.1	900	1 SYA_MYCPN	P75364 mycoplasma
20	87.5	4.1	880	1 GUN4_THERP	P26221 thermomios
21	87	4.1	999	1 MA2B_BOVIN	O24451 bos taurus
22	86	4.0	1066	1 SYL_PPRFO	O58792 pyrococcus
23	85.5	4.0	685	1 AMY1_DICFH	P09961 dictyoglomu
24	85.5	4.0	2405	1 DYHA_CHLHE	O36610 chlamydomon
25	85	4.0	339	1 LYCA_HOCHI	P15057 bacteriophia
26	85	4.0	700	1 CAN2_CHICK	O92178 gallus gall
27	85	4.0	833	1 CM41_YEAST	P53008 saccharomyc
28	84.5	3.9	570	1 RSL_CHDMT	P38016 chlamydia p
29	84.5	3.9	580	1 RSL_CHLUN	O59749 rhizobium p
30	84	3.9	758	1 PARC_RHIME	P21951 saccharomyc
31	84	3.9	2222	1 DPOF_YEAST	P05803 influenza a
32	83.5	4.9	470	1 NRAM_IAMIM	P13685 escherichia
33	83.5	4.9	472	1 PNB_ECOLI	

34	83.5	3.9	513	1 AR09_YEAST	P38640 saccharomyc
35	83.5	3.9	906	1 RPOD_BKRL1	P18147 bacterioph
36	83.5	3.9	1839	1 CYAA_SACKL	P23466 saccharomyc
37	83	3.9	293	1 RSO_NEUCR	O01291 neorospira
38	83	3.9	578	1 TRPA_RHMIT	P19813 cryptococcus
39	83	3.9	649	1 GUND_CLOTH	P04954 cistothidium
40	83	3.9	812	1 CHEA_TREPA	P96123 treponema p
41	83	3.9	884	1 RPOD_BPT3	P07459 bacterioph
42	83	3.9	901	1 MALC_ECOLI	P06993 escherichia
43	82.5	3.8	301	1 LEGL_MOUSE	P24721 mus musculu
44	82.5	3.8	489	1 TCRP_VIRCH	P29481 vitro to chlo
45	82.5	3.8	528	1 MDLC_PSEPO	P20906 pseudomonas

ALIGNMENT

RESULT	1	STANDARD	PRT	402 AA
RHP_PIG				
ID	RHP_PIG			
AC	P17560			
DT	01-AUG-1990 (Ref. 15, Created)			
DT	01-AUG-1990 (Ref. 15, Last sequence update)			
DT	30-MAY-2000 (Ref. 39, Last annotation update)			
DE	N-ACETYLGLUCOSAMINE 2-EPIMERASE (EC 5.1.3.8) (GLUCNAZ 2-EPIMERASE)			
DE	(N-ACETYL-D-GLUCOSAMINE 2-EPIMERASE) (RENIN-BINDING PROTEIN) (RHP)			
GN	RHP			
OS	Sus scrofa (pig)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 4-12 AND 326-357.			
RC	TISSUE=KIDNEY;			
KX	MEDLINE: 90216671.			
KA	Inoue H., Fukui K., Takahashi S., Miyake Y.;			
RT	"Molecular cloning and sequence analysis of a cDNA encoding a porcine			
RT	kidney renin-binding protein.";			
RL	J. Biol. Chem. 265:6556-6561(1990).			
RN	[2]			
RP	LEUCINE-ZIPPER, AND MOTACINISIS.			
KX	MEDLINE: 91268065.			
KA	Inoue H., Takahashi S., Fukui K., Miyake Y.;			
RT	"Leucine zipper motif in porcine renin-binding protein (RbP) and its			
RT	relationship to the formation of an RbP-renin heterodimer and an			
RT	RbP homodimer.";			
RL	J. Biol. Chem. 266:11896-11900(1991).			
RN	[3]			
RP	CHARACTERIZATION.			
RC	TISSUE=KIDNEY CORTEX;			
KX	MEDLINE: 96279179.			
KA	Maru I., Ohta Y., Murata K., Tsukada Y.;			
RT	"Molecular cloning and identification of N-acetyl-D-glucosamine			
RT	2-epimerase from porcine kidney as a renin-binding protein.";			
RL	J. Biol. Chem. 271:16294-16299(1996).			
CC	-1- FUNCTION: CATALYZES THE INTERCONVERSION OF N-ACETYLGLUCOSAMINE TO			
CC	N-ACETYLGLUCOSAMINE. BINDS TO RENIN FORMING A PROTEIN COMPLEX			
CC	CATALYZES HIGH MOLECULAR WEIGHT (HMW) RENIN AND INHIBITS RENIN			
CC	ACTIVITY.			
CC	-1- CATALYTIC ACTIVITY: N-ACETYL-D-GLUCOSAMINE - N-ACETYL-D-			
CC	GLUCNAZ 2-EPIMERASE ACTIVITY.			
CC	-1- PATHWAY: N-ACETYLGLUCOSAMINE ACID BIOSYNTHESIS.			
CC	-1- SUBUNIT: HOMODIMER.			
CC	-1- TISSUE SPECIFICITY: KIDNEY, LIVER, ADRENAL, AND PITUITARY GLANDS.			
CC	THE AMOUNT BEING MUCH GREATER IN KIDNEY THAN IN THE OTHER TISSUES.			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION			
CC	AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS			
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY			
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL			
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/anonymous/			

[illegible]

Environ Biol Fish (2015) 98:1039–1050
 Eikeland J, Mølleratun G, Jordal K, Taniuchi A, Verbeuren J, Entoslootm J, Eikeland J, Mølleratun G, Jordal K, Taniuchi A, Verbeuren J, Entoslootm J


```

11  SYV MY(TH)  STANI(AH);  876 AA.
12  05.4175
13

```

101 (b) MAY 2000 (Ref. 49, last annotation update)

IN VALS OR HV24H¹ OR MIV10H.04C.

Order 10. Filicinellales; Actinobacteroidales; Actinobacteroidales

I
I
I
N

K⁺ SIKKAIN H₂/H₂O:

VI

Holdstock K., Haslam D., Brown D., Hillinworth L., Connor K.,

Horvath T., Jurek K., Krogh A., McLellan J., Mouton S., Murphy

Kutler, Z., Skelton, S., Sindles, K., Silsion, J., and Quail, M.A., 1994, *Robots*, 2nd ed. (Cambridge, MA: Cambridge University Press).

Raymond S. Barrett, B.Sc.,
"Disciplined the biology of Mycobacterium tuberculosis from the
ket"

1. *Computer Graphics: Principles and Practice*, 2nd ed., by van Dam, van Dam, and van Dam. Addison-Wesley, 1995.

$$\text{PYRUVATE} + \text{VAL} + \text{TRNA(VAL)} \rightarrow \text{AMP}$$

SUBORDINATE MEMBER (BY SIMILARITY).

SIMILARITY: HETEROGENEOUS TO CLASS I AMINOACYL-TRNA SYNTHETASE FAMILY

This *SWISS* Photo gallery is copyright-free. It is produced through a collaboration between the Swiss Institute of Biotechnology and the EMBL outstation in the European Bioinformatics Institute. There are no restrictions on its use. By non-profit institutions as long as its content is in no way modified and this statement is not removed, disseminated for commercial purposes and this statement is not removed. (See: <http://www.isb.scrib.ch/announcements.html> for contact to license this site.)

[illegible]

Abstract

Figure 1

Diagram illustrating the experimental setup for measuring the effect of temperature on the rate of reaction between hydrogen peroxide and potassium iodide.

The diagram shows two test tubes labeled A and B, each containing a solution of hydrogen peroxide and potassium iodide. The test tubes are placed in water baths at different temperatures: Test tube A is in a cold water bath, and Test tube B is in a warm water bath. Arrows indicate the flow of gas (oxygen) being produced from the reaction in both test tubes. The gas is collected in inverted graduated cylinders over water. The volume of gas collected in cylinder B is visibly greater than in cylinder A, indicating a faster reaction rate at the higher temperature.

[illegible]

1. **INTRODUCTION**
 2. **THEORY**
 3. **EXPERIMENTAL**
 4. **CONCLUSION**
 5. **REFERENCES**
 6. **APPENDIX**
 7. **ACKNOWLEDGMENTS**
 8. **NOTES**
 9. **REFERENCES**
 10. **APPENDIX**
 11. **ACKNOWLEDGMENTS**
 12. **NOTES**
 13. **REFERENCES**
 14. **APPENDIX**
 15. **ACKNOWLEDGMENTS**
 16. **NOTES**
 17. **REFERENCES**
 18. **APPENDIX**
 19. **ACKNOWLEDGMENTS**
 20. **NOTES**
 21. **REFERENCES**
 22. **APPENDIX**
 23. **ACKNOWLEDGMENTS**
 24. **NOTES**
 25. **REFERENCES**
 26. **APPENDIX**
 27. **ACKNOWLEDGMENTS**
 28. **NOTES**
 29. **REFERENCES**
 30. **APPENDIX**
 31. **ACKNOWLEDGMENTS**
 32. **NOTES**
 33. **REFERENCES**
 34. **APPENDIX**
 35. **ACKNOWLEDGMENTS**
 36. **NOTES**
 37. **REFERENCES**
 38. **APPENDIX**
 39. **ACKNOWLEDGMENTS**
 40. **NOTES**
 41. **REFERENCES**
 42. **APPENDIX**
 43. **ACKNOWLEDGMENTS**
 44. **NOTES**
 45. **REFERENCES**
 46. **APPENDIX**
 47. **ACKNOWLEDGMENTS**
 48. **NOTES**
 49. **REFERENCES**
 50. **APPENDIX**
 51. **ACKNOWLEDGMENTS**
 52. **NOTES**
 53. **REFERENCES**
 54. **APPENDIX**
 55. **ACKNOWLEDGMENTS**
 56. **NOTES**
 57. **REFERENCES**
 58. **APPENDIX**
 59. **ACKNOWLEDGMENTS**
 60. **NOTES**
 61. **REFERENCES**
 62. **APPENDIX**
 63. **ACKNOWLEDGMENTS**
 64. **NOTES**
 65. **REFERENCES**
 66. **APPENDIX**
 67. **ACKNOWLEDGMENTS**
 68. **NOTES**
 69. **REFERENCES**
 70. **APPENDIX**
 71. **ACKNOWLEDGMENTS**
 72. **NOTES**
 73. **REFERENCES**
 74. **APPENDIX**
 75. **ACKNOWLEDGMENTS**
 76. **NOTES**
 77. **REFERENCES**
 78. **APPENDIX**
 79. **ACKNOWLEDGMENTS**
 80. **NOTES**
 81. **REFERENCES**
 82. **APPENDIX**
 83. **ACKNOWLEDGMENTS**
 84. **NOTES**
 85. **REFERENCES**
 86. **APPENDIX**
 87. **ACKNOWLEDGMENTS**
 88. **NOTES**
 89. **REFERENCES**
 90. **APPENDIX**
 91. **ACKNOWLEDGMENTS**
 92. **NOTES**
 93. **REFERENCES**
 94. **APPENDIX**
 95. **ACKNOWLEDGMENTS**
 96. **NOTES**
 97. **REFERENCES**
 98. **APPENDIX**
 99. **ACKNOWLEDGMENTS**
 100. **NOTES**
 101. **REFERENCES**
 102. **APPENDIX**
 103. **ACKNOWLEDGMENTS**
 104. **NOTES**
 105. **REFERENCES**
 106. **APPENDIX**
 107. **ACKNOWLEDGMENTS**
 108. **NOTES**
 109. **REFERENCES**
 110. **APPENDIX**
 111. **ACKNOWLEDGMENTS**
 112. **NOTES**
 113. **REFERENCES**
 114. **APPENDIX**
 115. **ACKNOWLEDGMENTS**
 116. **NOTES**
 117. **REFERENCES**
 118. **APPENDIX**
 119. **ACKNOWLEDGMENTS**
 120. **NOTES**
 121. **REFERENCES**
 122. **APPENDIX**
 123. **ACKNOWLEDGMENTS**
 124. **NOTES**
 125. **REFERENCES**
 126. **APPENDIX**
 127. **ACKNOWLEDGMENTS**
 128. **NOTES**
 129. **REFERENCES**
 130. **APPENDIX**
 131. **ACKNOWLEDGMENTS**
 132. **NOTES**
 133. **REFERENCES**
 134. **APPENDIX**
 135. **ACKNOWLEDGMENTS**
 136. **NOTES**
 137. **REFERENCES**
 138. **APPENDIX**
 139. **ACKNOWLEDGMENTS**
 140. **NOTES**
 141. **REFERENCES**
 142. **APPENDIX**
 143. **ACKNOWLEDGMENTS**
 144. **NOTES**
 145. **REFERENCES**
 146. **APPENDIX**
 147. **ACKNOWLEDGMENTS**
 148. **NOTES**
 149. **REFERENCES**
 150. **APPENDIX**
 151. **ACKNOWLEDGMENTS**
 152. **NOTES**
 153. **REFERENCES**
 154. **APPENDIX**
 155. **ACKNOWLEDGMENTS**
 156. **NOTES**
 157. **REFERENCES**
 158. **APPENDIX**
 159. **ACKNOWLEDGMENTS**
 160. **NOTES**
 161. **REFERENCES**
 162. **APPENDIX**
 163. **ACKNOWLEDGMENTS**
 164. **NOTES**
 165. **REFERENCES**
 166. **APPENDIX**
 167. **ACKNOWLEDGMENTS**
 168. **NOTES**
 169. **REFERENCES**
 170. **APPENDIX**
 171. **ACKNOWLEDGMENTS**
 172. **NOTES**
 173. **REFERENCES**
 174. **APPENDIX**
 175. **ACKNOWLEDGMENTS**
 176. **NOTES**
 177. **REFERENCES**
 178. **APPENDIX**
 179. **ACKNOWLEDGMENTS**
 180. **NOTES**
 181. **REFERENCES**
 182. **APPENDIX**
 183. **ACKNOWLEDGMENTS**
 184. **NOTES**
 185. **REFERENCES**
 186. **APPENDIX**
 187. **ACKNOWLEDGMENTS**
 188. **NOTES**
 189. **REFERENCES**
 190. **APPENDIX**
 191. **ACKNOWLEDGMENTS**
 192. **NOTES**
 193. **REFERENCES**
 194. **APPENDIX**
 195. **ACKNOWLEDGMENTS**
 196. **NOTES**
 197. **REFERENCES**
 198. **APPENDIX**
 199. **ACKNOWLEDGMENTS**
 200. **NOTES**
 201. **REFERENCES**
 202. **APPENDIX**
 203. **ACKNOWLEDGMENTS**
 204. **NOTES**
 205. **REFERENCES**
 206. **APPENDIX**
 207. **ACKNOWLEDGMENTS**
 208. **NOTES**
 209. **REFERENCES**
 210. **APPENDIX**
 211. **ACKNOWLEDGMENTS**
 212. **NOTES**
 213. **REFERENCES**
 214. **APPENDIX**
 215. **ACKNOWLEDGMENTS**
 216. **NOTES**
 217. **REFERENCES**
 218. **APPENDIX**
 219. **ACKNOWLEDGMENTS**
 220. **NOTES**
 221. **REFERENCES**
 222. **APPENDIX**
 223. **ACKNOWLEDGMENTS**
 224. **NOTES**
 225. **REFERENCES**
 226. **APPENDIX**
 227. **ACKNOWLEDGMENTS**
 228. **NOTES**
 229. **REFERENCES**
 230. **APPENDIX**
 231. **ACKNOWLEDGMENTS**
 232. **NOTES**
 233. **REFERENCES**
 234. **APPENDIX**
 235. **ACKNOWLEDGMENTS**
 236. **NOTES**
 237. **REFERENCES**
 238. **APPENDIX**
 239. **ACKNOWLEDGMENTS**
 240. **NOTES**

THE UNIVERSITY OF CHICAGO

DR	FRINIS ^a : F800096; 16ASYNTHAL.
FR	PROSITE: F800178; AA TRNA LIGASE 1; 1.
KW	AMINOACYL tRNA synthetase; tRNAse; ATP binding.
F1	SIMILAR 44 53 "HIGH" REGION
F1	SIMILAR 540 544 "KMSK" REGION
F1	SIMILAR 544 544 "ATP (HY STIMULATION)"
FT	RINITING: 876; AA: 97840 MM: 854E2C3EF7D4090 C604.
SD	SEQUENCE:

Only Match	4.46	Score	95	14	Length	6.76	
Post Local Similarity	19.86	Prod. No.	4.2				
Matches	74	Unsuccessful	47	Mismatches	141	Totals	167

	25	WREXSLDROGNOVETLOKROVOT	TERTIUM:	KOVANAVENITEF	78
09	111	11	1	1	1
10	117	WRKS	GZLQOMRICHYWSRIKFTIMOLISPAVETPKRYLATIY	MATR	171

07	140	172	196	220	244	268	292	316	340	364	388	412	436	460	484	508	532	556	580	604	628	652	676	700	724	748	772	796	820	844	868	892	916	940	964	988	1012	1036	1060	1084	1108	1132	1156	1180	1204	1228	1252	1276	1300	1324	1348	1372	1396	1420	1444	1468	1492	1516	1540	1564	1588	1612	1636	1660	1684	1708	1732	1756	1780	1804	1828	1852	1876	1900	1924	1948	1972	1996	2020	2044	2068	2092	2116	2140	2164	2188	2212	2236	2260	2284	2308	2332	2356	2380	2404	2428	2452	2476	2500	2524	2548	2572	2596	2620	2644	2668	2692	2716	2740	2764	2788	2812	2836	2860	2884	2908	2932	2956	2980	3004	3028	3052	3076	3100	3124	3148	3172	3196	3220	3244	3268	3292	3316	3340	3364	3388	3412	3436	3460	3484	3508	3532	3556	3580	3604	3628	3652	3676	3700	3724	3748	3772	3796	3820	3844	3868	3892	3916	3940	3964	3988	4012	4036	4060	4084	4108	4132	4156	4180	4204	4228	4252	4276	4300	4324	4348	4372	4396	4420	4444	4468	4492	4516	4540	4564	4588	4612	4636	4660	4684	4708	4732	4756	4780	4804	4828	4852	4876	4900	4924	4948	4972	4996	5020	5044	5068	5092	5116	5140	5164	5188	5212	5236	5260	5284	5308	5332	5356	5380	5404	5428	5452	5476	5500	5524	5548	5572	5596	5620	5644	5668	5692	5716	5740	5764	5788	5812	5836	5860	5884	5908	5932	5956	5980	6004	6028	6052	6076	6100	6124	6148	6172	6196	6220	6244	6268	6292	6316	6340	6364	6388	6412	6436	6460	6484	6508	6532	6556	6580	6604	6628	6652	6676	6700	6724	6748	6772	6796	6820	6844	6868	6892	6916	6940	6964	6988	7012	7036	7060	7084	7108	7132	7156	7180	7204	7228	7252	7276	7300	7324	7348	7372	7396	7420	7444	7468	7492	7516	7540	7564	7588	7612	7636	7660	7684	7708	7732	7756	7780	7804	7828	7852	7876	7900	7924	7948	7972	7996	8020	8044	8068	8092	8116	8140	8164	8188	8212	8236	8260	8284	8308	8332	8356	8380	8404	8428	8452	8476	8500	8524	8548	8572	8596	8620	8644	8668	8692	8716	8740	8764	8788	8812	8836	8860	8884	8908	8932	8956	8980	9004	9028	9052	9076	9100	9124	9148	9172	9196	9220	9244	9268	9292	9316	9340	9364	9388	9412	9436	9460	9484	9508	9532	9556	9580	9604	9628	9652	9676	9700	9724	9748	9772	9796	9820	9844	9868	9892	9916	9940	9964	9988	100
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

Lib	2.2/	AIAN	-----	HPDERYRULVGT	-----	SLAPVUJREAI	256
LY	190	LPPTVEVIACTVREVMOTDPLPPLGJMEAVNITGEVJNSE	-----	-----	-----	-----	244

[illegible]

294 PEOPLE WOULD...WON'T LIVE...-ALAKG...HOATGOKCMWNEPRNDYA... 538
 345 SEKSEPEPEPRISLOWWVRESLAKAANLAVRNBDTVLHPASMEKKNFSWYIMMHWIS 404

DY	3.9	- - - WSHFALPEYG	3.6
DH	4.05	ECLIMMCHKIPWYC	4.18

RESULT: 10
MAZE_HUMAN
(1) MAZE_HUMAN STANDARD: PRI: 1010 AA.
A: (00754; 074094; 016680; 015340;

PT 15 JUL-1998 (col. 36, Last sequence update)
 DT 05 MAY-2000 (col. 39, Last annotation update)
 DE LYSO-SMAL, ALPHA-MANNOSIDASE DEFECTORS (EC 4.2.1.24) (MANNOSEDASE;
 DE ALPHA-H) (LYSO-SMAL, ACID ALPHA-MANNOSIDASE) (LAMAN).

05 Homo sapiens (Human);
06 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
07 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
08 [1]

RA Nilsson G., Betz T., Rälise H., M.F., Ramchandran U., Eypen G., Hansen G.M., Malm D., Traub-Daatz A., Tachibana T., Tachibana T., K. MEDLINE: 97301768.

KL Alpha-mannosidosis: functional cloning of the lysosomal alpha-mannosidase cDNA and identification of a mutation in two affected siblings.¹²

KL Hum. Mol. Genet. 6: 717-726 (1997).

RE SEQUENCE FROM N.A.
EX MEDLINE: 97336044.
RA KILISE H.M.F., 2009, 1, NITSEN O., ROMERO G., TOLFOSTAD O.K.,
KA <http://doi.org/10.1155/2009/14723>

KT (MANH)_n
KL (MANH)_{4.2} : 2H₂O : 2H₂O (1997).

RN [3]
 RP SEQUENCE OF 23-1010 FROM N.A.
 RC TISSUE-SPLEEN;
 RA MEDLINE: 97067056.
 RA Lind Y.-F., Lai A., Moremen K.W.;
 RT "Cloning, expression, purification, and characterization of the human
 RT broad specificity lysosomal acid alpha-mannosidase.";
 RL J. Biol. Chem. 271:28348-28358(1996).
 RN [4]
 RP SEQUENCE OF 23-1010 FROM N.A.
 RA MEDLINE: 94220092.
 RA Nebes V.L., Schmidt M.C.;
 RT "Human lysosomal alpha-mannosidase: isolation and nucleotide sequence
 RT of the full-length cDNA.";
 RL Biochem. Biophys. Res. Commun. 200:239-245(1994).
 RN [5]
 RP PARTIAL SEQUENCE;
 RA MEDLINE: 95114211.
 RA Emiliadi G., Martiño S., Stirling J.L., Maras B., Orlicchio A.;
 RT "Partial sequence of the purified protein confirms the identity of
 RL the cDNA coding for human lysosomal alpha-mannosidase B.";
 RL Biochem. J. 305:363-366(1995).
 RN [6]
 RP VARIANTS AM LEU-71: ARG-355 AND TRP-749.
 RA MEDLINE: 98411785.
 RA Gotoh Y., Wakamatsu N., Kawai H., Nishida Y., Matsumoto T.;
 RT "Missense and nonsense mutations in the lysosomal alpha-mannosidase
 RT gene (MANB) in severe and mild forms of alpha-mannosidosis.";
 RL Am. J. Hum. Genet. 63:1015-1024(1998).
 RN [7]
 RP VARIANTS AM, AND VARIANTS.
 RA MEDLINE: 99115084.
 RA Berg I., Risse H.M.F., Hansen G.M., Malm D., Tranobjaerg I.;
 RA Jørgensen O.K., Nilsen O.;
 RT "Spectrum of mutations in alpha-mannosidosis.";
 RL Am. J. Hum. Genet. 64:77-88(1999).
 RN [8]
 RP FUNCTION: NECESSARY FOR THE CATABOLISM OF N-LINKED CARBOHYDRATES
 RELEASED DURING GLYCOPROTEIN TURNOVER. CLEAVES ALL KNOWN TYPES OF
 ALPHA-MANNOSIDIC LINKAGES.
 RN [9]
 RP CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING
 ALPHA-D-MANNOSE RESIDUES IN ALPHA-D-MANNOSIDES.
 RN [10]
 RP SUBCELLULAR LOCATION: LYSOSOMAL.
 RN [11]
 RP PTM: FIRST PROCESSED INTO 3 PEPTIDES OF 70 KDA, 42 KDA (D) AND
 14/15 KDA (E). THE 70 KDA PEPTIDE IS FURTHER PROCESSED INTO THREE
 PEPTIDES (A, B AND C). THE A, B AND C PEPTIDES ARE
 DISULFIDE-LINKED.
 RN [12]
 RP PTM: HEAVILY GLYCOSYLATED.
 RN [13]
 RP DISEASE: DEFECTS IN MANB ARE THE CAUSE OF LYSOSOMAL ALPHA-
 MANNOSIDOSIS (AM), A LYSOSOMAL STORAGE DISEASE CHARACTERIZED BY
 ACCUMULATION OF UNBRANCHED OLIGOSACCHARIDES CHAINS. THIS
 VACUOLATION PREDOMINANTLY IN THE CNS AND PARENCHYMATOUS ORGANS,
 DEPENDING ON THE CLINICAL FINDINGS AT THE AGE OF ONSET, A SEVERE
 INFANTILE (TYPE I) AND A MILD JUVENILE (TYPE II) FORM OF ALPHA-
 MANNOSIDOSIS ARE RECOGNIZED. THERE IS CONSIDERABLE VARIATION IN
 THE CLINICAL EXPRESSION WITH MENTAL RETARDATION, RECURRENT
 INFECTIONS, IMPAIRED HEARING AND HURLER-LIKE SKELETAL CHANGES
 BEING THE MOST CONSISTENT ABNORMALITIES.
 RN [14]
 RP SIMILARITY: BELONGS TO FAMILY 38 OF GLYCOSYL HYDROLASES.
 RN [15]
 RP This SWISS-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL: 060266; AAC34130.1; -;
 DR EMBL: 060899; AAC51362.1; -;
 DR EMBL: 060885; AAC51362.1; JOINED.
 DR EMBL: 060886; AAC51362.1; JOINED.
 DR EMBL: 060887; AAC51362.1; JOINED.

DR EMBL: 060888; AAC51362.1; JOINED.
 DR EMBL: 060889; AAC51362.1; JOINED.
 DR EMBL: 060890; AAC51362.1; JOINED.
 DR EMBL: 060891; AAC51362.1; JOINED.
 DR EMBL: 060892; AAC51362.1; JOINED.
 DR EMBL: 060893; AAC51362.1; JOINED.
 DR EMBL: 060894; AAC51362.1; JOINED.
 DR EMBL: 060895; AAC51362.1; JOINED.
 DR EMBL: 060896; AAC51362.1; JOINED.
 DR EMBL: 060897; AAC51362.1; JOINED.
 DR EMBL: 060898; AAC51362.1; JOINED.
 DR EMBL: 068567; AAC50812.1; -;
 DR EMBL: 005572; AAB03816.1; -;
 DR MIM: 248500; -;
 DR INTERPRO: IPR000602; -;
 DR PFAM: PF01074; Glycosyl hydrolase 38; 1.
 KW Hydrolase; Glycosylase; Glycoprotein; Lysosome; Zymogen; Signal;
 KW Disease mutation; Polymorphism.
 RN [16]
 RP SIGNAL
 FT CHAIN 1 48
 FT CHAIN 49 1010
 FT CHAIN 49 344
 FT CHAIN 345 428
 FT CHAIN 429 600
 FT CHAIN 601 881
 FT CHAIN 882 1010
 FT CARBOHYD 132 132
 FT CARBOHYD 309 309
 FT CARBOHYD 366 366
 FT CARBOHYD 496 496
 FT CARBOHYD 644 644
 FT CARBOHYD 650 650
 FT CARBOHYD 691 691
 FT CARBOHYD 765 765
 FT CARBOHYD 831 831
 FT CARBOHYD 929 929
 FT CARBOHYD 988 988
 FT CARBOHYD 71 71
 FT VARIANT 277 277
 FT VARIANT 311 311
 FT VARIANT 336 336
 FT VARIANT 354 354
 FT VARIANT 355 355
 FT VARIANT 401 401
 FT VARIANT 412 412
 FT VARIANT 713 713
 FT VARIANT 749 749
 FT VARIANT 808 808
 FT CONFLICT 185 185
 FT CONFLICT 342 342
 FT CONFLICT 383 383
 SQ SEQUENCE 1010 AA; 113673 MW; 11425F6A340F3BE1 CkCk4;
 Query Match 4.48; Score 94.5; DR 1; Length 1010;
 Best Local Similarity 17.7%; Pred. No. 4.2;
 Matches 71; Conservative 46; Mismatches 134; Indels 151; Gaps 17;
 OY 89 ARHRDDQWNYFAIDGCKRIPKOPVNVPSDCFAAMAFSSQYALASQAFARALAIQAVN 147
 DB 11 AKGGLDSACHWIMS-----RAIDKPIPLPI--CPFLILAAACAGAGYETCPVQNMN 64
 OY 148 -NVLRRQIHPKG-----QY-----EKSYGRPLKSLAVELIANTL-----L 184

06	64	WILLFHHHOMWIKLVYUFGYGLKNDIGHUAVUJLLESLVSAALDHPFVYLAFF	128
07	185	EXEMHLLPTVEEVLACVYKREVMHDLTLEHGLMKREAVTQRELVNSF	256
08	124	SHWMDUQINATQFVAVKOLVROKLEFANQVMMHKAATYGAVWQMLDRLPHQ	184
09	233	IKR-----LNP-----GHEMMQMMQADSRKQJQFQAVLNTLE	278
10	184	NHJGKRVAMHLLPHTSHSFGASLFLACMGTGTFE-----GHLQYQKQVVRKQKLEME	255
07	275	YAW-----PHEHSHLPYFLDKQJHQQJLMD-----	401
16	236	QVWPKSTSLKJLADLELGV-----LNNYNNPKNDVWLVQJULVEHPSTLENAKEIV	292
07	402	-----GRKMWLELLVNAKSHGATQVE-----	425
10	293	IAFELNVALAGREYVRLNHLVMTDMSDQYENANMWFENLDRLLKLVANQAQKQSSVHAY	428
09	426	KJQWQFQVHUYAWK-----HFAQJYQJGMPYTLNR	457
10	453	SLFATYLMELAKRANLQWCVKRLDFTYVADHLDJWTFYFSKR	494

[illegible]

FT	DOMAIN	2402	2493	2494	2495	2496	2497	2498	2499	2500	2501	2502	2503	2504	2505	2506	2507	2508	2509	2510	2511	2512	2513	2514	2515	2516	2517	2518	2519	2520	2521	2522	2523	2524	2525	2526	2527	2528	2529	2530	2531	2532	2533	2534	2535	2536	2537	2538	2539	2540	2541	2542	2543	2544	2545	2546	2547	2548	2549	2550	2551	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561	2562	2563	2564	2565	2566	2567	2568	2569	2570	2571	2572	2573	2574	2575	2576	2577	2578	2579	2580	2581	2582	2583	2584	2585	2586	2587	2588	2589	2590	2591	2592	2593	2594	2595	2596	2597	2598	2599	2600	2601	2602	2603	2604	2605	2606	2607	2608	2609	2610	2611	2612	2613	2614	2615	2616	2617	2618	2619	2620	2621	2622	2623	2624	2625	2626	2627	2628	2629	2630	2631	2632	2633	2634	2635	2636	2637	2638	2639	2640	2641	2642	2643	2644	2645	2646	2647	2648	2649	2650	2651	2652	2653	2654	2655	2656	2657	2658	2659	2660	2661	2662	2663	2664	2665	2666	2667	2668	2669	2670	2671	2672	2673	2674	2675	2676	2677	2678	2679	2680	2681	2682	2683	2684	2685	2686	2687	2688	2689	2690	2691	2692	2693	2694	2695	2696	2697	2698	2699	2700	2701	2702	2703	2704	2705	2706	2707	2708	2709	2710	2711	2712	2713	2714	2715	2716	2717	2718	2719	2720	2721	2722	2723	2724	2725	2726	2727	2728	2729	2730	2731	2732	2733	2734	2735	2736	2737	2738	2739	2740	2741	2742	2743	2744	2745	2746	2747	2748	2749	2750	2751	2752	2753	2754	2755	2756	2757	2758	2759	2760	2761	2762	2763	2764	2765	2766	2767	2768	2769	2770	2771	2772	2773	2774	2775	2776	2777	2778	2779	2780	2781	2782	2783	2784	2785	2786	2787	2788	2789	2790	2791	2792	2793	2794	2795	2796	2797	2798	2799	2800	2801	2802	2803	2804	2805	2806	2807	2808	2809	2810	2811	2812	2813	2814	2815	2816	2817	2818	2819	2820	2821	2822	2823	2824	2825	2826	2827	2828	2829	2830	2831	2832	2833	2834	2835	2836	2837	2838	2839	2840	2841	2842	2843	2844	2845	2846	2847	2848	2849	2850	2851	2852	2853	2854	2855	2856	2857	2858	2859	2860	2861	2862	2863	2864	2865	2866	2867	2868	2869	2870	2871	2872	2873	2874	2875	2876	2877	2878	2879	2880	2881	2882	2883	2884	2885	2886	2887	2888	2889	2890	2891	2892	2893	2894	2895	2896	2897	2898	2899	2900	2901	2902	2903	2904	2905	2906	2907	2908	2909	2910	2911	2912	2913	2914	2915	2916	2917	2918	2919	2920	2921	2922	2923	2924	2925	2926	2927	2928	2929	2930	2931	2932	2933	2934	2935	2936	2937	2938	2939	2940	2941	2942	2943	2944	2945	2946	2947	2948	2949	2950	2951	2952	2953	2954	2955	2956	2957	2958	2959	2960	2961	2962	2963	2964	2965	2966	2967	2968	2969	2970	2971	2972	2973	2974	2975	2976	2977	2978	2979	2980	2981	2982	2983	2984	2985	2986	2987	2988	2989	2990	2991	2992	2993	2994	2995	2996	2997	2998	2999	3000																																																																																																																
SD	SEQUENCE	2493	AA:	2719/79	MM:	106AB/2760	FFR:	64.4																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						

10	10	HEAVY CHAIN	STANDARD	PROT.	657	AA.
11	11	PIG CHAIN	STANDARD	PROT.	657	AA.
12	12	PIG CHAIN	STANDARD	PROT.	657	AA.
13	13	PIG CHAIN	STANDARD	PROT.	657	AA.
14	14	PIG CHAIN	STANDARD	PROT.	657	AA.
15	15	PIG CHAIN	STANDARD	PROT.	657	AA.
16	16	PIG CHAIN	STANDARD	PROT.	657	AA.
17	17	PIG CHAIN	STANDARD	PROT.	657	AA.
18	18	PIG CHAIN	STANDARD	PROT.	657	AA.
19	19	PIG CHAIN	STANDARD	PROT.	657	AA.
20	20	PIG CHAIN	STANDARD	PROT.	657	AA.
21	21	PIG CHAIN	STANDARD	PROT.	657	AA.
22	22	PIG CHAIN	STANDARD	PROT.	657	AA.
23	23	PIG CHAIN	STANDARD	PROT.	657	AA.
24	24	PIG CHAIN	STANDARD	PROT.	657	AA.
25	25	PIG CHAIN	STANDARD	PROT.	657	AA.
26	26	PIG CHAIN	STANDARD	PROT.	657	AA.
27	27	PIG CHAIN	STANDARD	PROT.	657	AA.
28	28	PIG CHAIN	STANDARD	PROT.	657	AA.
29	29	PIG CHAIN	STANDARD	PROT.	657	AA.
30	30	PIG CHAIN	STANDARD	PROT.	657	AA.
31	31	PIG CHAIN	STANDARD	PROT.	657	AA.
32	32	PIG CHAIN	STANDARD	PROT.	657	AA.
33	33	PIG CHAIN	STANDARD	PROT.	657	AA.
34	34	PIG CHAIN	STANDARD	PROT.	657	AA.
35	35	PIG CHAIN	STANDARD	PROT.	657	AA.
36	36	PIG CHAIN	STANDARD	PROT.	657	AA.
37	37	PIG CHAIN	STANDARD	PROT.	657	AA.
38	38	PIG CHAIN	STANDARD	PROT.	657	AA.
39	39	PIG CHAIN	STANDARD	PROT.	657	AA.
40	40	PIG CHAIN	STANDARD	PROT.	657	AA.
41	41	PIG CHAIN	STANDARD	PROT.	657	AA.
42	42	PIG CHAIN	STANDARD	PROT.	657	AA.
43	43	PIG CHAIN	STANDARD	PROT.	657	AA.
44	44	PIG CHAIN	STANDARD	PROT.	657	AA.
45	45	PIG CHAIN	STANDARD	PROT.	657	AA.
46	46	PIG CHAIN	STANDARD	PROT.	657	AA.
47	47	PIG CHAIN	STANDARD	PROT.	657	AA.
48	48	PIG CHAIN	STANDARD	PROT.	657	AA.
49	49	PIG CHAIN	STANDARD	PROT.	657	AA.
50	50	PIG CHAIN	STANDARD	PROT.	657	AA.
51	51	PIG CHAIN	STANDARD	PROT.	657	AA.
52	52	PIG CHAIN	STANDARD	PROT.	657	AA.
53	53	PIG CHAIN	STANDARD	PROT.	657	AA.
54	54	PIG CHAIN	STANDARD	PROT.	657	AA.
55	55	PIG CHAIN	STANDARD	PROT.	657	AA.
56	56	PIG CHAIN	STANDARD	PROT.	657	AA.
57	57	PIG CHAIN	STANDARD	PROT.	657	AA.
58	58	PIG CHAIN	STANDARD	PROT.	657	AA.
59	59	PIG CHAIN	STANDARD	PROT.	657	AA.
60	60	PIG CHAIN	STANDARD	PROT.	657	AA.
61	61	PIG CHAIN	STANDARD	PROT.	657	AA.
62	62	PIG CHAIN	STANDARD	PROT.	657	AA.
63	63	PIG CHAIN	STANDARD	PROT.	657	AA.
64	64	PIG CHAIN	STANDARD	PROT.	657	AA.
65	65	PIG CHAIN	STANDARD	PROT.	657	AA.
66	66	PIG CHAIN	STANDARD	PROT.	657	AA.
67	67	PIG CHAIN	STANDARD	PROT.	657	AA.
68	68	PIG CHAIN	STANDARD	PROT.	657	AA.
69	69	PIG CHAIN	STANDARD	PROT.	657	AA.
70	70	PIG CHAIN	STANDARD	PROT.	657	AA.
71	71	PIG CHAIN	STANDARD	PROT.	657	AA.

DR	PDB:	1HC6; 4I-JAN-94.
DR	PDB:	1HC6; 4I-JAN-94.
DR	PDB:	1HC7; 4I-JAN-94.
DR	INTERPRO:	IPRO00896; -
DR	INELPROP:	IPRO0227; -
DR	PFAM:	PF00372; hemoCyalin_1.
DR	PRINTS:	PR00187; HEMOCYALIN.
DR	PROSITE:	PS00209; HEMOCYANIN_1.
DR	PROSITE:	PS00210; HEMOCYANIN_2.
DR	PROSITE:	PS00498; CYCNOUSASE_2; 1.
KW	Respiratory protein; oxygen transport; copper; glycoprotein;	
KW	Betalymph; xD-structure.	
ET	DISTILED	483 98
ET	DISTILED	562 502
F1	CARBOHYD	167 609
F1	METAL	194 194
FT	METAL	198 198
ET	METAL	224 224
ET	METAL	444 444
F1	METAL	448 448
F1	METAL	484 484
F1	TORN	8 9
ET	HELIX	10 17
ET	TORN	18 18
ET	TORN	21 22
F1	HELIX	28 36
F1	TORN	39 40
F1	TORN	43 44
FT	STRAND	46 46
ET	TORN	47 48
FT	TORN	51 52
F1	HELIX	53 61
F1	TORN	62 63
F1	TORN	74 75
ET	HELIX	77 91
ET	TORN	92 92
ET	STRAND	93 93
F1	HELIX	96 105
F1	TORN	106 109
ET	HELIX	112 125
ET	HELIX	127 129
ET	HELIX	137 140
ET	HELIX	142 144
F1	TORN	145 145
F1	HELIX	148 160
F1	TORN	161 161
ET	STRAND	166 167
ET	TORN	174 176
ET	TORN	179 180
F1	HELIX	182 185
F1	TORN	186 186
F1	HELIX	188 200
ET	TORN	203 204
ET	HELIX	207 209
ET	TORN	216 217
F1	HELIX	218 236
F1	TORN	237 239
F1	TORN	248 249
ET	STRAND	251 252
ET	STRAND	261 261
ET	STRAND	267 267
F1	STRAND	275 276
F1	STRAND	280 280
F1	TORN	281 283
ET	STRAND	284 284
ET	TORN	286 286
ET	HELIX	287 402
F1	STRAND	404 406
F1	TORN	408 409
F1	STRAND	412 414
ET	TORN	418 419
ET	HELIX	420 428

ET	TORN	4.29	4.80
ET	TORN	3.82	3.83
ET	TORN	3.37	3.39
ET	TORN	3.41	3.42
ET	HELIX	3.43	3.52
ET	TORN	3.54	3.55
ET	HELIX	3.67	3.69
ET	TORN	3.71	3.73
ET	HELIX	3.71	3.76
ET	TORN	3.78	3.78
ET	HELIX	3.79	3.95
ET	TORN	3.96	3.96
ET	HELIX	4.03	4.06
ET	STRAND	4.12	4.19
ET	STRAND	4.23	4.26
ET	STRAND	4.29	4.31
ET	HELIX	4.33	4.36
ET	STRAND	4.50	4.52
ET	STRAND	4.55	4.57
ET	STRAND	4.61	4.67
ET	STRAND	4.74	4.83
ET	TORN	4.97	4.98
ET	HELIX	4.99	5.01
ET	STRAND	5.03	5.12
ET	STRAND	5.17	5.22
ET	TORN	5.23	5.25
ET	TORN	5.27	5.28
ET	STRAND	5.29	5.30
ET	TORN	5.36	5.37
ET	HELIX	5.38	5.44
ET	TORN	5.45	5.49
ET	TORN	5.56	5.57
ET	STRAND	5.64	5.65
ET	HELIX	5.66	5.68
ET	STRAND	5.78	5.88
ET	HELIX	5.89	5.92
ET	TORN	5.94	5.93
ET	TORN	6.11	6.12
ET	TORN	6.22	6.25
ET	TORN	6.33	6.35
ET	HELIX	6.36	6.38
ET	TORN	6.40	6.41
ET	STRAND	6.42	6.51
NO	SEQUENCE	6.57 AA:	7.56-95 MM: 1E5520E0426671085116004:

```

Query Match 4.28; Score 91; DB 1; Length 457;
Best Local Similarity 22.58; Prod. No. 4.7;
Matches 64; Conservative 35; Mismatches 96; Indels 86; Gaps 17;

QY 18 HQVVLFWFKYSIDKQSGGMYPTLQKQKGVPLDQKFMQNRGV--WQFAVYNRLELRK 75
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 198 HMDFFFWWHSY-----GYH--LQKQEL-----FFWVHQGLAKPLDPRISNNMLD 244
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 76 Q--WELELR-----HAGDPLAR-----HGRVQDG-----NMYE 101
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 ELHMRLTREDQAPRLSYKYGGDFVPRPDNIHPEIVDVAHYNDLELLESTHEALDHY 304
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 102 ALDQGRK--PLRQVYV--FSLQCFAMAPQVYALASCAQFAKALQGVNNVLRQGNHR 157
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 ITDSGHTITLQKQKQELTSDI EESSKYSNNVQYGSLLNTA-----HVMLSRQSGDH 358
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 158 GQY-----FKSYDGIKQSLAVMIILANILEMMLDPLVEVLAQVKEVMIDF 210
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 359 GKNLPQVMIHPELAKQDSFFRLHKIMQNFKKHIDSTPRYHNT-----FF 408
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 211 LDPQLMFEAVTPGE---FVDSERGLIN--FQHGTE 244
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 409 ----SCWVNVIALDDELFLPFDQVQYSLINAVDSQENLE 444
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Job: 1 time: 460 3000

GenBank version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein: protein search, using sw model

Run on: January 1, 2001, 22:50:15 : Search time 74.18 Seconds

(without alignments)
492.172 Million cell updates/sec

Title: US-09-645-321-1

Perfect score: 2147

Sequence: 1 MIAHRQELAAQVYVGAHQD.....CFIVPRALMLAEILQEPS 391

Scoring table: HUSUM62

Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 9437416 residues

Total number of hits satisfying chosen parameters: 297973

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPITREML_14:
1: SP_archaea:
2: SP_bacteria:
3: SP_fungi:
4: SP_human:
5: SP_invertebrate:
6: SP_mammal:
7: SP_mirc:
8: SP_oranicle:
9: SP_phage:
10: SP_plant:
11: SP_protist:
12: SP_virus:
13: SP_vertebrate:
14: SP_unclassified:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2147	100.0	491	2	P74124
2	691	32.2	492	6	Q95331
3	662	30.8	419	11	P82343
4	108.5	5.1	607	2	P74605
5	103	4.8	662	5	Q20180
6	98.5	4.5	422	10	Q95185
7	96.5	4.5	758	10	Q95185
8	94.5	4.4	266	5	Q45352
9	92	4.3	1067	1	Q9V072
10	91.5	4.3	692	2	Q67902
11	91.5	4.3	996	10	Q97VW2
12	91	4.2	991	10	Q97VW2
13	90.5	4.2	554	5	Q97916
14	90.5	4.2	880	3	Q97880
15	90	4.2	417	2	Q95910
16	90	4.2	506	8	Q91K88
17	89.5	4.2	662	4	Q960X1
18	89.5	4.2	673	4	Q960X2
19	89.5	4.2	2932	5	Q9V075

20	88.5	4.1	658	2	Q92477
21	88.5	4.1	698	3	Q04563
22	88.5	4.1	801	3	Q74344
23	88.5	4.1	1161	11	Q9V077
24	88	4.1	469	5	Q9V0W2
25	88	4.1	113	5	Q9V0V6
26	87.5	4.1	1248	5	Q9V0B6
27	87	4.1	348	2	Q9V0P9
28	87	4.1	375	2	Q9S5K0
29	87	4.1	536	2	Q9ZNP4
30	87	4.1	720	3	Q59838
31	86.5	4.0	642	1	Q28537
32	86	4.0	787	13	Q9V0V4
33	85.5	4.0	344	9	Q9ZK42
34	85.5	4.0	723	3	Q94632
35	85.5	4.0	912	1	Q9V0E0
36	85	4.0	266	5	Q90543
37	85	4.0	349	9	Q38009
38	84.5	3.9	511	4	Q16422
39	84.5	3.9	583	13	Q9V0T4
40	84.5	3.9	729	13	Q91147
41	84.5	3.9	731	13	Q91150
42	84	3.9	465	2	Q05508
43	84	3.9	646	5	Q9V114
44	83.5	3.9	546	2	Q9R870
45	83.5	3.9	760	10	Q9S805

ALIGNMENTS

RESULT	1	PRELIMINARY:	PR1:	491 AA:
ID	P74124			
AC	P74124:			
DT	01-FEB-1997 (TEMBLrel. 02, Created)			
PT	01-FEB-1997 (TEMBLrel. 02, Last sequence update)			
DE	01-JUN-2000 (TEMBLrel. 14, Last annotation update)			
DE	RENIN-BINDING PROTEIN.			
GN	SLR1975.			
OS	Synochrysis sp. (strain PCC 6803).			
OC	Bacteria; Cyanobacteria; Chlorococcales; Synochrysis.			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE: 97061201.			
RA	Kaneko T., Sato S., Kotani H., Tanaka A., Asanuma E., Nakamura Y.,			
RA	Myajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,			
RA	Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okamura S.,			
RA	Shimpo S., Takeuchi C., Wada T., Watanabe A., Yanada M.,			
RA	Tabata S.;			
RT	*Sequence analysis of the genome of the unicellular cyanobacterium			
RT	Synochrysis sp. strain PCC6803. 11. Sequence determination of the			
RT	entire genome and assignment of potential protein-coding regions. "			
RL	DNA Res. 3:109-136(1996).			
DR	EMBL: D90912; IMA18210.1; "			
SV	SEQUENCE: 491 AA; 45630 MW; C386F4AD0BEFA140 CRC64;			
Query Match	100.0%	Score 2147:	DB 2:	Length 491:
Best Local Similarity	100.0%	Prod. No. 4,781,776:		
Matches 391:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	1	MIAHRQELAAQVYVGAHQDVLPRWKSILDSAGVFTCDKRGVFTDKIPLWGNK	60	
DB	1	MIAHRQELAAQVYVGAHQDVLPRWKSILDSAGVFTCDKRGVFTDKIPLWGNK	60	
QY	61	WQFAVFNRLPEPKWELIARHAEFLARHGRDQNNMYFALDQESKPLKQYNNVSR	120	
DB	61	WQFAVFNRLPEPKWELIARHAEFLARHGRDQNNMYFALDQESKPLKQYNNVSR	120	
QY	121	FAAMASVAYALASACAKAATACQAVNNVLRQNNPKQVYKSYGTRIPKSLAVPMIA	180	
DB	121	FAAMASVAYALASACAKAATACQAVNNVLRQNNPKQVYKSYGTRIPKSLAVPMIA	180	

```

RESULT 2
ID 096341 PRELIMINARY: PRT: 402 AA.
AC 096341:
DT 01 FEB 1997 (ITERATED: 02, created)
DT 01 FEB 1997 (ITERATED: 02, last sequence update)
DT 01 JAN 1999 (ITERATED: 09, last annotation update)
DE N-ACETYL D GLUCOSAMINE 2 EPIMERASE (EC 5.1.3.8).
OS Mus musculus (Mouse).
OC Embryonal: Choroida; Craniala; Vertebrata; Euteleostomi;
OC Mammalia: Eutheria; Cetartiodactyla; Suidae; Suidae Sus.
KN [1]
KA SEQUENCE FROM N.A.
KA PLATZER M., BREUER V., KOTHEWALD K., WIEDE L., OKSCHKE A.,
KA "Comparative sequence analysis of the mouse Irfam locus and the
KA corresponding region of human K926."
KT Submitted (MAR 1999) to the EMBL/GenBank/DDBJ databases.
CC 1. FUNCTION: CATALYZES THE INTERCONVERSION OF N-ACETYLGLUCOSAMINE TO
CC N-ACETYLMANNOSAMINE. BINDS TO KENIN FORMING A PROTEIN COMPLEX
CC CALLED HIGH MOLECULAR WEIGHT (HMW) KENIN AND INHIBITS KENIN
CC ACTIVITY (BY SIMILARITY).
CC 1. CATALYTIC ACTIVITY: N-ACETYL D GLUCOSAMINE + N-ACETYL D
CC MANNOSAMINE.
CC 1. ENZYME REGULATION: ACTIVITY ENHANCED BY ATP (BY SIMILARITY).
CC 1. PATHWAY: N-ACETYLHEXAMINIC ACID BIOSYNTHESIS.
CC 1. SUBUNIT: HOMODIMER (BY SIMILARITY).
DR EMBL: AF134093; AAF22157.1;
DR MGI: MGI:105940; RndpP.
KW Isoenzyme.
LT DOMAIN.
SO SPODINNE 402 AA; 402 MW; 409146760471606 CIRC64;

```

```

Quality Match 40.8% Score 662; DB 11; Length 419;
Host Local Similarity 45.5% Prod. No. 8.5e 49;
Matches 140; Conserved 75; Mismatches 154; Indels 48; Gaps 8;
5 PROGLAQQYVALHQD...VLPWEKYSIDRGSGYFTGLDRKQGVPLDRLKIMQNGV 61
3 KERETLQWKKRVEDELDQVLAEMWESHSDGEGFTGLDRKQGVPLDRLKIMQNGV 62
62 WQFAVYNGLE...FKQWLELADHDLADLADNR...DQGNMYVALDQERKLRQYVNF 117
118 STCFAMAFSQVALASAGAKAIALQAVNNVLR...KQHNPKQDYKESYGTPLKSLAVP 176
123 SECTYTMANNEILKRVKGVGEVYQSEALFMMDGILLPWKREPALDNGVLSGALATPRMVP 182
177 MLANLILM...EMLDPTVEEVALQVREVMIDDELDELGLMRA 221
183 MMLSLVQLGEGDEDELIMNYALQW...CVNILLQHVGR...DEQVAL...EN 228
222 VPTGGEVVSERKLLNPGIGLEAMFMMDIAGKSDRGDLOGLAVLAVLNTLEVMRDE 280
229 VSRHEELNCTQKHONPGLDQVLAQVLAQKRDQKLMNIDKPELIDPHSGMDQE 288
281 FQSLFEYLDGSGHFGDQVLMQDKIMWVHLETLVALAKGQATQGRK...WQMFERYHIVAMS 440
289 IRSLFEYQVADQVLTQLEEMNKILMWPISBAMIAELMYSISGSDALLDLFEKVAAYTFR 448
441 IRADPEYQWPMVYINRQGVLLNKGKRWK...FENVALKMAFETLQGLAVS 491
449 QFRPEYQWPMVYINRQGVLLNKGKRWK...FENVALKMAFETLQGLAVS 499

```



```

10 17 4606 PRE: MINARY: PRT: 607 AA.
AC 17 4606:
11 01-FEB-1997 (TREMURIEL, 02, created)
12 01-FEB-1997 (TREMURIEL, 02, last sequence update)
13 01-JUN-2000 (TREMURIEL, 14, last annotation update)
14 HYPOTHEICAL: 70.0 KDa PROTEIN.
15 SIK1855.
16 Synchocystis sp. (strain pcc 6803).
17 Bacteria: Cyanobacteria: Chroococcales: Synchocystis.
18 [1]
19 SEQUENCE FROM N.A.
20 MEDLINE: 97061201.
21 Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
22 Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
23 Hoshino T., Murano A., Muraki A., Nakazaki N., Nario K., Okumura S.,
24 Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
25 Takata S.
26 "Sequence analysis of the genome of the unicellular cyanobacterium
27 Synchocystis sp. strain pcc6803. 1. Sequence determination of the
28 entire genome and assignment of potential protein-coding regions."
29 J. Mol. Biol. 310:1-16(1996).
30 EMBL: D90508; BAI17650.1; -.
31 Hypothetical protein.
32 KW: SEQUENCE: 607 AA; 70003 MW; 873E42F37B6A0F3 CRC64;
33
34 Query Match: 5.1%; Score 108.5; DB 2; Length 607;
35 Best Local Similarity: 22.7%; Pred. No. 0.42; Indels 129; Gaps 25;
36 Matches 89; Conservative 47; Mismatches 127;
37
38 58 NQVQVAFVFNKRLPPKQWLEIARHGAQPIARHGR--TQNGN---WYFALVQDQKPLKQ 112
39 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
40 159 SLVYGFALF--MHTGQVLEAEKGEYELREHREYKEDIIYVHALDVQGEKQK 217
41 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
42
43 114 EV--NFSIE--FAAM--AFSQ--YALASQAEKATALQVNNVLRQ--IN----- 155
44 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
45 218 FASFGVGVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 269
46 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
47
48 156 -FKQYKSYKQTRPL-----KSLA-----VMTLANTLLEKMLPPT 194
49 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 270 LKSEYGFTHSLDPLMLPSESLGPNKARKNNNSVADHAPVYLNL-----WL--AT 321
51 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
52
53 195 VEEVLA-----QVREVMTPLEPDIOMEAVPTGEV-----DSFGRL 237
54 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
55 422 GQKYADMLETFDTTEKFPDY-----DISPFQDERFEDMSIDITWQWQ 368
56 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57
58 208 N---FPHGLEAMWEMQIAVRSQDQLOEOALAT--LNTLEAVDEEFGITFYLDK--QG 292
59 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
60 409 NAAVVSNNIKIAWNIKRNKSLKSKRYVLAARKIADIMAVDSQDQDQWVAVVERLLIN 428
61 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62
63 294 HRP--QULEW--DQKLMWVLETLVA--LA-----KRIQATNDEKQWQFEVHYVAMSIF 342
64 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 429 HSKCHQFVWHKRAKWDQDALVLLLAGITIDEEYHRIHQE----- 471
66 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67
68 343 ADPEYQWPGVILNRQ---EVLLN---LKGQ 367
69 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70 472 ASAEVYAMETLDEGGITENVLANIPYLAG 504
71 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72
73 RESULT 5
74 Q95H40 PRELIMINARY: PRT: 662 AA.
75
76 01-NOV-1996 (TREMURIEL, 01, created)
77 01-NOV-1996 (TREMURIEL, 01, last sequence update)
78 01-MAY-2000 (TREMURIEL, 13, last annotation update)
79 HEMOCYANIN PRECURSOR [CONTAINS: HEMOCYANIN, LONG FORM, HEMOCYANIN,
80 SHORT FORM].
81 Penaeus vannamei (Penaeid shrimp) (European white shrimp).
82 Eukaryota: Metazoa: Arthropoda: Crustacea: Malacostraca:
83 Eumalacostraca: Eucarida: Decapoda: Penillopteroidea: Penaeoidea:
84 Penaeidae: Penaeus.
85 [1]

```

```

86 SEQUENCE FROM N.A. AND SEQUENCE OF 15-52.
87 TISSUE=HEPATOXYANIN.
88 MEDLINE: 97409475.
89 Seelios P.Y., Lemoine S., Van Wormhout A.:
90 "Molecular cloning of hemocyanin cDNA from Penaeus vannamei
91 (Crustacea, Decapoda): structure, evolution and physiological
92 aspects."
93 FEBS Lett. 407:15-18(1997).
94
95 1- FUNCTION: HEMOCYANIN ARE COPPER CONTAINING OXYGEN CARRIERS.
96 OCCURRING FREELY DISSOLVED IN THE HEMOLYMPH OF MANY MOLLUSKS AND
97 ARTHROPODS.
98 EMBL: X82502; CAA57880.1; -.
99 HSSP: P04254; 1H71.
100 INTERPRO: IPR000846; -.
101 INTERPRO: IPR002227; -.
102 PFAM: PF00472; hemocyanin_1.
103 PRINTS: PR00187; HEMOCYANIN.
104 PROSITE: PS00209; HEMOCYANIN_1; 1.
105 PROSITE: PS00210; HEMOCYANIN_2; 1.
106 PROSITE: PS00498; LYCOSTINASE_2; 1.
107 KW: Signal: Respiratory protein; oxygen transport; copper;
108 Endoplasmic reticulum; Metal-binding.
109 FT SIGNAL: 1 14
110 FT CHAIN: 15 662 HEMOCYANIN, LONG FORM.
111 FT PROPEP: 15 662
112 FT CHAIN: 21 662 HEMOCYANIN, SHORT FORM.
113 FT METAL: 208 208 COPPER A (BY SIMILARITY).
114 FT METAL: 212 212 COPPER A (BY SIMILARITY).
115 FT METAL: 236 236 COPPER B (BY SIMILARITY).
116 FT METAL: 358 358 COPPER B (BY SIMILARITY).
117 FT METAL: 362 362 COPPER B (BY SIMILARITY).
118 FT METAL: 398 398 COPPER B (BY SIMILARITY).
119 SEQUENCE: 662 AA; 74980 MW; 52B0137557E75178 CRC64;
120
121 Query Match: 4.8%; Score 104; DB 5; Length 662;
122 Best Local Similarity: 22.7%; Pred. No. 1.4;
123 Matches 87; Conservative 47; Mismatches 149; Indels 104; Gaps 21;
124
125 23 PPKRYSLKGGGTYFTGKRGQVETDKFTLQNGNQGVAF--FYNGLEP--KPYQL 78
126 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 216 PEW-----WNAVGHHLQKQ-----ENFWLHQLTVFPAERLSNYLQVGHQWN 263
128 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129
130 79 E-----IARHGAPLARHG-----KQNGWVYFALQ 105
131 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
132
133 264 KPIYDIAPIPIITYKQGVGFARPIQNVKFEVDVARTREAVYESKRIATAIHYT--VDS 322
134 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
135
136 106 FQKPL---KQYVNVSDCTAAMASQYALASQAEKATATQAVNNVLRQNGKQVY 161
137 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 323 EKKHIDISNKGIDIDITDITSSLYSPNVQYGCALHNIA-----HIVDQDGHQKFD 376
139 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140
141 162 KSYPS-----TRPLKSLAVPMLNLTLEMMLPPT-----VEEVATVYRE 205
142 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 377 LP--PQVLEHFFATIRIDISPRDIAKRYMNIIPKHNIDIPVTKAIDPESVSLIYAVGE 435
144 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145
146 206 VMTDIPDPEICIMEAVPTGEFVNSFEGRLNQHGLEAMWEMQIAVRSQDQLOEOAL 265
147 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 436 LFTYFETFEYSLIN--AVTQAEGLTEWEISTYVRLNKEFTITLIVENGAER----- 487
149 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
150
151 324 QEKQWQFEVHYVAMSIF--FAQPEYQ 348
152 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
153 533 EKKSITSSVIVDVDSIDHIFAFAPAG 559
154 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155
156 RESULT 6
157 Q95H5 PRELIMINARY: PRT: 422 AA.
158
159 01-MAY-2000 (TREMURIEL, 14, created)

```


DE SERINE THREONINE PROTEIN KINASE SNF1P.
 GN SNF1.
 GS Cochlodolus carduum (Bipolaris zeicola).
 GA Cochlodolus carduum SNF1 gene is required for cell wall degradation
 GA Eukaryota: Fungi: Ascomycota: Pleosporales: Pleosporaceae:
 GA Cochlodolus.
 KA STRAIN SH11.
 KA STRAIN SH11.
 KA Tooukari N.J., Scott-Craig J.S., Walton J.D.:
 KA "Cochlodolus carduum SNF1 gene is required for cell wall degradation
 KA enzyme expression and virulence on maize."
 KA Submitted (JUN 1999) to the EMBL/GenBank/DBJ databases.
 KA EMBL: AF159253; ADB3441.1; .
 KA INTERPRO: IPR000719; .
 KA INTERPRO: IPR002290; .
 KA EFM: PF00069; PKinase_1.
 KA PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
 KA PROSITE: PS0108; PROTEIN_KINASE_ATP_1.
 KA PROSITE: PS0011; PROTEIN_KINASE_P006; 1.
 KA Kinase.
 KA SEQUENCE 880 AA; 97949 MW; CDIDG124624E16 CRC64;
 Query Match 4.28; Score 90.57; DB 3; Length 880;
 Best Local Similarity 19.18; Pred. No. 25;
 Matches 49; Conservative 35; Mismatches 102; Indels 71; Gaps 9;
 QY 25 WEKSLLEKQKQVFTLEKRGVFTDKFVLQKQVQVAFVFNRIEPRKQMLEIRH; 84
 DB 564 WPKVIRKRAK;-----KSSS-----KSSISQ-----APRRKRSRHS 618
 QY 85 ADFIARHQRKQWYVATLQDQKPLKQPVNSFKFAAMASQYALASQAFKATALQ 144
 DB 619 QSTSSHSSEDDG-----SRKGSREPLSVRN-----GSTQELKRGK; 661
 QY 145 AVNVLKQKQKQYKSTYQTRKSLAVPMLANIILPMKLLPTTIVEVIAQIVR 204
 DB 662 HYNH-----NMQVYVQEDVWVTHARPL-----KQMPFPGVAHNSSTHSR 704
 QY 205 EVMLEFLPELIMEAVPTGFEVDFEGRLLNPGHTEAMWPMQIAQSGRQLQEQ 264
 DB 705 ---VHANSQARRKRSSTNIS-----TSAGHGVKRMILSPKAGVSSEIHVNIDE 752
 QY 265 AIAVVLNILEYAWDEP 281
 DB 753 AVYVMSIQLYSIDRF 769
 RESULT 15
 Q59910 PRELIMINARY: PRT: 417 AA.
 A: Q59910;
 D: 01 NOV 1996 (FIREHUTEL_01, Created)
 D: 01 NOV 1996 (FIREHUTEL_01, Last sequence update)
 D: 01 MAY 2000 (FIREHUTEL_13, Last annotation update)
 D: CYTOCHROME P450 (F02.1.14.14.1).
 OS Streptomyces filialiae.
 OT Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
 OT Actinomycetales: Streptomycesineae: Streptomycetaceae: Streptomyces.
 KA SEQUENCE FROM N.A.
 KA STRAIN SH11.
 KA STRAIN-159253.
 KA MEDLINE: 95075319.
 KA Meisner Davies L.A., Condliffe E.:
 KA "Analysis of liver tylosin biosynthetic genes from the tylosin region of
 KA the Streptomyces fradiae genome."
 KA Mol. Microbiol. 13:339-355(1994).
 OS SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DE EMBL: D08223; AAA21341.1; .
 DE HSSP: G00441; 10XA.
 DE INTERPRO: IPR001128; .
 DE INTERPRO: IPR002497; .
 DE Pfam: PF00067; P450; 1.

DB PRINIS: P00059; HP450.
 DB PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN 1.
 KW oxidoreductase; Monooxygenase; Electron transport; Membrane; Iron-
 FT BINDING 466 466 HEME (BY SIMILARITY).
 SU SEQUENCE 417 AA; 47224 MW; A76194224DFCB1 CRC64;

Query Match 4.28; Score 90; DB 2; Length 417;
 Best Local Similarity 21.28; Pred. No. 9.4;
 Matches 77; Conservative 36; Mismatches 111; Indels 140; Gaps 20;
 QY 55 WLDNQVQWQ-----FAVNYNLEKPKQWFIARHGAAP 148HSD 94
 DB 50 WDSKQWQVRYDYDVLVSINQDPSDSQVFPVPELALIMQDIFQGIHDEH; 107
 QY 95 QDQNWYFALLQEKPLKQPVNSDFPFAMAFSQYALASQAFKATALQAVNNVIRKH 154
 DB 108 -----PLRK-----LVSAAPPRKATLEPRATLEPQLDQ----- 140
 QY 155 NPKQYKSYNGTRPKSLAVNMLA-----NLTLMKML-----LP--PIVFEV 198
 DB 141 REKGQIDVQSLAVPLPVIAELLGIPAEPRDLEKRWVIVLNNQMEYFNIPQHS 200
 QY 199 LAQIVKRWMLDPLDQIMREAVPTGFEVDFEGRLLNPGHTEAMWPMQIAQSGR 240
 DB 201 MCPALKEWQVLYRRLAKRE--TLDLMSGLLPAVHGRRLDEELVNIVALILLAG 257
 QY 241 HQLFA-----MWMQW-----IQKSTDRQLQEQALAVNILEYAWDEPSEITFFL 288
 DB 258 HISSATILQNLPLVLDHREDAELRADRIIDGATE--ELLRYK--SPFNNTFLRKE 312
 QY 289 --DROGHPPQW--LEMDQKLMWVHLFTVALAKHQATIQKQKQWQFERYVDAWSP 342
 DB 313 DTDLLGIPKAKQGVVAV-----IASANDSA-----HF 341
 QY 343 ADPE 346
 DB 342 SDPD 345

Search completed: January 1, 2001, 22:57:47
 Job time: 452 sec

[illegible]

ALLEGMENTS

RESULT	1
PROBAND	196091.2
DEFINITION	1,6-OTUS
ACCESSION	160091.2
VERSION	160091.2.1
DESCRIPTION	1,6-OTUS, bp, DNA, 807P, 07-APR-1999 SPOXYNUCLEIN ACID, 5'-phosphoryl-5'-phosphoryl, 14-27, 1719644-1846241. 160091.2.1: G1:1654228 4'-ethoxydeoxyribose, 4'-ethoxydeoxyribose, 40S ribosomal protein S1: 40S

WILLIAM
J. HILL

REFLECTION

Appendix

Kamigaito, N.; Sawada, S.; Kojima, H.; Tanaka, A.; Asanuma, E.; Nakamura, Y.
Kamigaito, N.; Hirose, M.; Saito, M.; Sato, M.; Sato, S.; Kimura, J.;
Hosokoshi, T.; Matsuno, A.; Morita, A.; Nakazaki, N.; Naito, K.;
Kamigaito, N.; Shimizu, S.; Takahashi, T.; Wada, T.; Watanabe, A.;
Yamada, M.; Yasuda, M.; and Ishida, Y.

TITLE	Synuclein analysis of the genome of the microalga <i>Chlorella</i>
JOURNAL	<i>Synuclein</i>
MEDLINE	97061201
REFERENCE	2 (bases 1 to 128598)
AUTHORS	Tabata S.
TITLE	Direct Sequencing
JOURNAL	Journal of Molecular Biology

COMMENT
FEATURES
Potential protein coding regions were assigned on the basis of similarity search of the ORFs and GeneMark analysis.
Location/Qualifications

7285
 /Estadillo "Pocobello" 114.4"
 /Ch. 81.1" - 114.0" : 114.4"
 /Campamento (1 - 343)

```

/codon_start=1
/insert_table=1
/product="hypothetical protein"
/protein_id="BAAL1144.1"
/db_xref="cd:1653229"
/description="MARIADYAVI,SGAALATMCKLQAVVATGNGVLTSDAEST
VNIASIVAWASIAATPAIDSAHPQHSKEVEYSSCEAGFVFAVATIDYSAVERI
LSPVATLQNALATATATATACGTAAVILMRASRLNLSIKMSQHMVTVISV
VSVAVALLVYGMELDELALAGVAVVWVLTGTLLETTISDMQSGSPVALQATV
CTELYEGVGRFHLLOTROASQSESTIRIVVPIHWVYGRHDLFALEIYAIABETI
SKVILFELIDKRSWQHPDEFPASPIINRKN"
comp_bcount(997,11442)

```

```

//seed_start=1
//trans_label=1
//product "hypothetical protein"
//protein_id="AA18145.1"
//db_xref="ref:1654230"
//translation="MPLPTLANKRPTKALENDQALAVVAIRKSTGGRYQRIKAK
GVASTSAQGLREVEALMDVIRIVRILKRIKIAQGVAVPTVFAVLTGSGYLENI
FAQSKIRLWMLIRYILSQTELDLSTIKRVRKIVLEMRIRHVRWQDILLK"
//completeness(1707...2363)

```

```

compElement (1707, -2363)
/quote-"151"
/quote-"00F-ID:5111261"
/cooId-start-1
/trans1-table-1
/product-"elongation factor 1S"
/protein_id-"BAAL18146.1"
/db_xref-"GI:1658231"
/translation-"MAETFDLVYKELKELTANMMKKAALKEKFTLEKFTMLKQK
GIASLKKSKHTAAGLEVSTIIRSGAGVLEVENLEIDVAFKQKRIKLVNIVAMQ
AATPNVAVSADIPQEVAAKKELEEMQKQKQKRIKRIKLVGRDKKIKELSLI
DQPTKQKNTLEELVQALAEADQNTGVRRFTLPFNQNTLEKAEINAEELVAAAKK

```

[illegible]

[illegible][illegible]

Query Match	12.28	Score 142.8	DB 88	Length 1292
Best Local Similarity	48.78	Prod. No. 110312		
Method	545	Conservative	0	Match 59.27
			Indels	21
			Gaps	5
polyA signal	1278			
polyA-site	1292			
HAStE-coatIN	299 a	327 c	376 g	290 t
ORIGIN				
Query Match	12.28	Score 142.8	DB 88	Length 1292
Best Local Similarity	48.78	Prod. No. 110312		
Method	545	Conservative	0	Match 59.27
			Indels	21
			Gaps	5
polyA signal	1278			
polyA-site	1292			
HAStE-coatIN	299 a	327 c	376 g	290 t
ORIGIN				

[illegible]

Query Match	Best Local Similarity	Score	ID#	Length
Matches: 540; Conservation: 0; Mismatches: 538; Indels: 21; Gaps: 52	48.3%	134.2	DB 3	1372
60				
140				
120				
190				
180				
250				
244				
310				
288				
370				
348				
440				
408				
490				
465				
550				
525				
610				
582				
670				
642				
730				
702				

[illegible][illegible]


```

repeat_region
  complement(3242, ..3593)
  /rpt_family="Alu-Jo"
  /rpt_family="Alu-Jo"
  complement(4509, ..4821)
  /rpt_family="Alu-Sp"
  5413, ..5697
repeat_region
  /rpt_family="Alu-Jo"
  complement(6126, ..6295)
  /rpt_family="Alu-J"
  complement(7050, ..7310)
  /rpt_family="Alu-Jb"
  complement(7402, ..7599)
  /rpt_family="Alu-Sp"
  8253, ..8529
repeat_region
  complement(9747, ..9926)
  /rpt_family="Alu-Jb"
  10090, ..10194
  /rpt_family="Alu-J"
  11157, ..11364
  /rpt_family="L1MB3"
  11593, ..11890
  /rpt_family="L1MC2"
  12083, ..12370
  /rpt_family="Alu-Sx"
  13165, ..13386
  /rpt_family="Alu-Jo"
  complement(13727, ..13784)
  /rpt_family="Alu-Sx2"
  complement(13801, ..13948)
  /rpt_family="L1MC2"
  13951, ..14005
  /rpt_family="Alu-J"
  complement(14468, ..15598)
  /gene="L1CAM"
  number=28
  complement(join(15599, ..15830, 16304, ..16315, 16413, ..16485,
  16819, ..16953, 17258, ..17413, 17521, ..17640, 17757, ..17940,
  18024, ..18146, 18235, ..18436, 18640, ..18755, 19585, ..19607,
  19900, ..20060, 20291, ..20488, 20735, ..20845, 20933, ..21057,
  21237, ..21393, 21496, ..21662, 21776, ..21887, 22455, ..22598,
  22738, ..22869, 22991, ..23175, 23323, ..23434, 23725, ..23895,
  23992, ..24114, 25090, ..25292, 25530, ..25635, 26166, ..26180,
  26699, ..26793))
  /gene="L1CAM"
  complement(14468, ..28793)
  /gene="L1CAM"
  complement(14468, ..28793)
  /gene="L1CAM"
  complement(14868, ..15830)
  /gene="L1CAM"
  number=28
  complement(join(15599, ..15830, 16304, ..16315, 16413, ..16485,
  16819, ..16953, 17258, ..17413, 17521, ..17640, 17757, ..17940,
  18024, ..18146, 18235, ..18436, 18640, ..18755, 19585, ..19607,
  19900, ..20060, 20291, ..20488, 20735, ..20845, 20933, ..21057,
  21237, ..21393, 21496, ..21662, 21776, ..21887, 22455, ..22598,
  22738, ..22869, 22991, ..23175, 23323, ..23434, 23725, ..23895,
  23992, ..24114, 25090, ..25292, 25530, ..25635, 26166, ..26180,
  26699, ..26793))
  /gene="L1CAM"
  /odon_start=1
  /odon_end=1
  /product="neutral cell adhesion molecule 11"
  /db_xref="NCI:130265d"
  /db_xref="NCI:130265d"
  /translat ion="MVAALRYVWPLLSPLLIQIPEEYGHMEPEVITEESPR
  LEVPEPDI SLKTEASGRPEVQEFWRGVHKEKRELSAVVQSHPSSFTITGNS
  NAEORFQIVGECASNKLTAMSHETRLMAGAPKREKTEKPEVEEESVLPNP
  PPSAETIKTYMNSKLIETKIDKAVIMQONINYPAAVLISSNSDIYCHAFGTRI
  LUKEPIDUKKATNSMIDKRPILPPTNSSSHVAJUGQPVLEETIARSPPTIKM
  LRSQPMADRVLYONHKLIOJLKVSEIDJGEPYCAJQNSISARHAYVIVAAAY
  WLKPSDILYGEETARLLVAVSRPEEVYWRINLPEVELADQYRIQRIALLS
  WUCSDIVITGCEARNRIGLLVNAVITYVWIKITACQUTYMAVQYSTAYILLKA
  FGAVPSVQWIDEDETTVIDEREFPANETIGIRLOANDTFRYCLANDONNVI
  MANIKVNIATOTIGQKSTETKRSKSVTPVQASPDSTQSTIWRGQRIQIOLHS
  DKETEDKRIYINSLYSIDQNTSCVASTELQVESHAQJLVQSPGVRIYISDILH

```

```

exon
  complement(16413, ..16485)
  /gene="L1CAM"
  number=26
  complement(16819, ..16953)
  /gene="L1CAM"
  number=25
  complement(17258, ..17413)
  /gene="L1CAM"
  number=24
  complement(17521, ..17640)
  /gene="L1CAM"
  number=23
  complement(17757, ..17930)
  /gene="L1CAM"
  number=22
  complement(18024, ..18146)
  /gene="L1CAM"
  number=21
  complement(18235, ..18436)
  /gene="L1CAM"
  number=20
  complement(18640, ..18755)
  /gene="L1CAM"
  number=19
  complement(19256, ..19418, ..19501)
  /rpt_family="Alu-Sx2"
  /rpt_family="MER42"
  complement(19585, ..19807)
  /gene="L1CAM"
  number=18
  complement(19900, ..20060)
  /gene="L1CAM"
  number=17
  complement(20291, ..20488)
  /gene="L1CAM"
  number=16
  complement(20735, ..20845)
  /gene="L1CAM"
  number=15
  complement(20933, ..21057)
  /gene="L1CAM"
  number=14
  complement(21237, ..21393)
  /gene="L1CAM"
  number=13
  complement(21496, ..21662)
  /gene="L1CAM"
  number=12
  complement(21776, ..21887)
  /gene="L1CAM"
  number=11
  complement(22455, ..22598)
  /gene="L1CAM"
  number=10
  complement(22738, ..22869)
  /gene="L1CAM"
  number=9
  complement(22991, ..23175)

```

```

LILUSQVHVSWSPIAGHNPALDKYDITFEUKIMALKWYSLAKVHINQISLILKUSIV
VHYIPEVYAIKNVGHQSDPSVSEIVVIEPAAEKEMVIVKQDINTIMVYIWKILPM
MNAVAQVYVYQVQMRPQGTREPMQDQVSDPILVSNISTEVEYELKVQANSQKRP
EPQVITIGVSGEDYQAIPELETIELINSSAVYKRPVYLAQVKGHLGAVYTWKRE
SQKRSKRITIKHDIIVVPANTTSVILSLIRYSSHLVQAINRQSGPASEETSTP
EAVYSDTAHLIEQNSISLILRWQPSHNVVIGYVISTHUTDESKQDLSINRQ
PEIKIHNILDISPHIKYRQDQATIKQGRHAIYHNSQIMASISQDNIISAIQEN
YVSVSWYKQSDQNEPRLIFKALCHEKQASISQVYVSNQSYFQWUQDQIDYEI
HILKEPHEQMAVKNSTGQVPELPASEATEGMEIIEVSVAIILILVILITIKRS
KIKKQSVKIKEDTOVISEARPKDETFEYRSELESNEERAKGSSQSPINIKRELS
IDISLAVGSSVQVQREDSITIGQYSGKREKDAQNSINQATSPINAVALE"

```


87117	87216	gap of	100 bp	in length
87217	88245	contig of	1029 bp	in length
88246	88434	gap of	100 bp	in length
88434	89493	contig of	1148 bp	in length
89494	89593	gap of	100 bp	in length
89594	90596	contig of	1003 bp	in length
90597	90656	gap of	100 bp	in length
90697	93150	contig of	2654 bp	in length
93351	93450	gap of	100 bp	in length
93451	94536	contig of	1086 bp	in length
94537	94636	gap of	100 bp	in length
94637	95797	contig of	1161 bp	in length
95798	95897	gap of	100 bp	in length
95898	97063	contig of	1166 bp	in length
97064	97163	gap of	100 bp	in length
97164	98244	contig of	1071 bp	in length
98245	98344	gap of	100 bp	in length
98345	99650	contig of	1316 bp	in length
99651	99750	gap of	100 bp	in length
99751	101345	contig of	1595 bp	in length
101346	101445	gap of	100 bp	in length
101446	107114	contig of	5669 bp	in length
107115	107214	gap of	100 bp	in length
107215	110997	contig of	3783 bp	in length
110998	111097	gap of	100 bp	in length
111098	114007	contig of	2910 bp	in length
114008	114107	gap of	100 bp	in length
114108	115183	contig of	1076 bp	in length
115184	115283	gap of	100 bp	in length
115284	118707	contig of	4424 bp	in length
118708	118807	gap of	100 bp	in length
118808	120082	contig of	1275 bp	in length
120083	120182	gap of	100 bp	in length
120183	122094	contig of	1912 bp	in length
122095	122194	gap of	100 bp	in length
122195	123403	contig of	1209 bp	in length
123404	123503	gap of	100 bp	in length
123504	127244	contig of	3741 bp	in length
127245	127344	gap of	100 bp	in length
127345	128970	contig of	1626 bp	in length
128971	129070	gap of	100 bp	in length
129071	130490	contig of	1420 bp	in length
130491	130590	gap of	100 bp	in length
130591	132146	contig of	1556 bp	in length
132147	132244	gap of	100 bp	in length
132245	137509	contig of	5263 bp	in length
137510	137609	gap of	100 bp	in length
137610	138922	contig of	1313 bp	in length
138923	139022	gap of	100 bp	in length
139023	141189	contig of	2167 bp	in length
141190	141289	gap of	100 bp	in length
141290	142374	contig of	1085 bp	in length
142375	142474	gap of	100 bp	in length
142475	144249	contig of	1775 bp	in length
144250	144349	gap of	100 bp	in length
144350	147088	contig of	2739 bp	in length
147089	147188	gap of	100 bp	in length
147189	148659	contig of	1471 bp	in length

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-128P23"
/clone_1db="RP11.1"
1..2308
misc_feature
/note="assembly_fragment:003336"
fragment_chain:1"
misc_feature
/note="assembly_fragment:01005"
fragment_chain:1"
4676..7354
/note="assembly_fragment:00638"

```

	Query Match	Best Local Similarity	3.4%	Score 49.8	DB 84:	Length 148579
	Matches 71:	Conservative 0:	Mismatches 52:	Indels 0:	Gaps 0:	
misc_feature	fragment_chain:1"	7455..8720	/note="assembly_fragment:00376	fragment_chain:1"	8821..10650	fragment="assembly_fragment:00514
misc_feature	fragment_chain:2"	10751..12192				
misc_feature						
QY 1030	gattccgagatgaggaatcgattcgactgaatcgccgagagagatgattactcaac	1089				
DB 110214	gttcggagagatcttgtaattgttcagatgtttgtatggatggcagagagtgtaaacgct	110274				
QY 1090	ctaaagagagagaaatgagaaagatgcttcctccgacactctgagatctatgaa	1149				
DB 110274	caactacgcctgcgaagctaacgcagagaaacacagctctccactcattctgactcaatctctg	110444				
QY 1150	aaa	1152				
DB 110334	aaa	110336				

RESULT	11
AL163539	
LOCUS	AL163539 154112 bp DNA HTG 14-JUN-2000
DEFINITION	Homo sapiens chromosome 13 clone RP11-444C16, *** SEQUENCING IN PROGRESS. ***
ACCESSION	AL163539
VERSION	AL163539.4 GI:9230998
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Ikayama; Metazoan; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eularchia; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 154112)
AUTHORS	Burton,J.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-2000) Sanger Centre, Hinxton, Cambridge, UK; CH10 ISA, UK. E-mail enquiries: hinxton@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk
COMMENT	On Jul 16, 2000 this sequence version replaced gi:8573850.

```

Center : saenger_centre
Center code: SC
Web site: http://www.saenger.ac.uk
Contact: humpu@saenger.ac.uk
-----
Project Information
Center project name: ba14401b
-----
Summary Statistics
Assembly program: XSeal4, version 4.5
Cheministry vector: plasmid; 108752; 100% of reads
Consensus quality: 131709 bases at least Q40
Consensus quality: 138219 bases at least Q40
Consensus quality: 142824 bases at least Q20
Insert size: 150212; sum-of-contigs
Insert size: 159639; aaarose-tp
Quality coverage: 2.85x in Q20 bases; sum-of-counts quality
coverage: 2.66x in Q20 bases; aaarose-tp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
4426: contig of 4426 bp in length
1

```

```

* 4427 4526: gap of 100 bp
* 4527 6166: count iq of 2649 bp in length
* 6167 6266: gap of 100 bp
* 6267 12956: count iq of 6690 bp in length
* 12957 14056: gap of 100 bp
* 14057 48191: count iq of 34145 bp in length
* 48192 48291: gap of 100 bp
* 48292 58146: count iq of 9845 bp in length
* 58147 68246: gap of 100 bp
* 68247 60162: count iq of 1926 bp in length
* 60163 62224: gap of 100 bp
* 62225 62425: count iq of 1962 bp in length
* 62426 64498: gap of 100 bp
* 64499 64598: count iq of 2174 bp in length
* 64599 70055: count iq of 5457 bp in length
* 70056 70155: gap of 100 bp
* 70156 74584: count iq of 4429 bp in length
* 74585 74684: gap of 100 bp
* 74685 77947: count iq of 3263 bp in length
* 77948 78047: gap of 100 bp
* 78048 79781: count iq of 1744 bp in length
* 79782 79881: gap of 100 bp
* 79882 80969: count iq of 1088 bp in length
* 80970 81069: gap of 100 bp
* 81070 84623: count iq of 3554 bp in length
* 84624 84723: gap of 100 bp
* 84724 85940: count iq of 1207 bp in length
* 85941 86040: gap of 100 bp
* 86041 87122: count iq of 1092 bp in length
* 87123 87222: gap of 100 bp
* 87223 89098: count iq of 1876 bp in length
* 89099 89198: gap of 100 bp
* 89199 91945: count iq of 2747 bp in length
* 91946 92045: gap of 100 bp
* 92046 94451: count iq of 1316 bp in length
* 94452 94551: gap of 100 bp
* 94552 94651: count iq of 1200 bp in length
* 94652 94751: gap of 100 bp
* 94752 95763: count iq of 1012 bp in length
* 95764 95863: gap of 100 bp
* 95864 97009: count iq of 1146 bp in length
* 97010 97109: gap of 100 bp
* 97110 99457: count iq of 2348 bp in length
* 99458 99557: gap of 100 bp
* 99558 102566: count iq of 3009 bp in length
* 102567 102666: gap of 100 bp
* 102667 109403: count iq of 6647 bp in length
* 109404 109403: gap of 100 bp
* 109404 110703: count iq of 1400 bp in length
* 110704 110803: gap of 100 bp
* 110804 113476: count iq of 2574 bp in length
* 113477 114476: gap of 100 bp
* 114477 118046: count iq of 4570 bp in length
* 118047 118146: gap of 100 bp
* 118147 119846: count iq of 1630 bp in length
* 119847 119946: gap of 100 bp
* 119947 123719: count iq of 3784 bp in length
* 123720 123819: gap of 100 bp
* 123820 128041: count iq of 4222 bp in length
* 128042 128141: gap of 100 bp
* 128142 129496: count iq of 1255 bp in length
* 129497 129496: gap of 100 bp
* 129497 131959: count iq of 2464 bp in length
* 131960 132059: gap of 100 bp
* 132060 138747: count iq of 6678 bp in length
* 138748 138847: gap of 100 bp
* 138848 141491: count iq of 2554 bp in length
* 141492 141491: gap of 100 bp
* 141492 142943: count iq of 1452 bp in length
* 142944 143043: gap of 100 bp
* 143044 147504: count iq of 4461 bp in length
* 147505 147604: gap of 100 bp

```

FEATURES

SOURCE

```

* 147605 149060: count iq of 1456 bp in length
* 149061 149160: gap of 100 bp
* 149161 151795: count iq of 2634 bp in length
* 151796 151895: gap of 100 bp
* 151896 154112: count iq of 2217 bp in length
Location/Qualifiers
1. 154112
  /organism "Homo sapiens"
  /db_xref "taxon:9606"
  /chromosome "13"
  /clone "RP11-344316"
  /clone_1fb "RP11.2"
1. 3426
  /note "assembly: fragment:00163"
  /fragment_chain:1"
4527. 6166
  /note "assembly: fragment:00424"
  /fragment_chain:1
  clone_end:590
  vector_side:left"
6267. 12956
  /note "assembly: fragment:01120"
  /fragment_chain:1"
14057. 48191
  /note "assembly: fragment:01176"
  /fragment_chain:2"
48292. 58146
  /note "assembly: fragment:00176"
  /fragment_chain:2"
58247. 60162
  /note "assembly: fragment:00615"
  /fragment_chain:2"
60263. 62224
  /note "assembly: fragment:00044"
  /fragment_chain:3"
62426. 64498
  /note "assembly: fragment:00845"
  /fragment_chain:4"
64599. 70055
  /note "assembly: fragment:00055"
  /fragment_chain:5"
70056. 70155
  /note "assembly: fragment:00116"
  /fragment_chain:6"
70156. 74584
  /note "assembly: fragment:00675"
  /fragment_chain:4"
74585. 77947
  /note "assembly: fragment:00925"
  /fragment_chain:5"
77948. 79781
  /note "assembly: fragment:00898"
  /fragment_chain:6"
79782. 80969
  /note "assembly: fragment:01043"
  /fragment_chain:6"
80970. 81069
  /note "assembly: fragment:00206"
  /fragment_chain:6"
81070. 84623
  /note "assembly: fragment:00142"
  /fragment_chain:7"
84624. 84723
  /note "assembly: fragment:00474"
  /fragment_chain:7"
84724. 85940
  /note "assembly: fragment:00124"
  /fragment_chain:8"
85941. 86040
  /note "assembly: fragment:00094"
  /fragment_chain:8"
86041. 87122
  /note "assembly: fragment:00114"
  /fragment_chain:8"
94752. 95763

```

misc_feature

	misc_feature	/note="assembly_traitment:00185"				
		95864..97009				
		/note-"assembly_fragment:00203"				
	misc_feature	97110..99457				
		/note-"assembly_fragment:00300"				
	misc_feature	99558..102566				
		/note-"assembly_trament:00464"				
	misc_feature	102667..109303				
		/note-"assembly_framgent:00479"				
Oxy Match:						
Best Local Similarity	57.7%	Pred.No. 1.9;				
Matches	Conservative	0; Mismatches	52; Indels	0; Gaps	0;	
OY	1030 gatacctaattatggaagatttgcttgcgtacctgatcatcccgaggagaagtactaacac	1089				
Ib	122946 GTTCTGGGAGCATGTGTAATTCACCGTGTTCTGTCCGCCGCACCAGAGTGCACAACCC	122995				
OY	1090 CTAAAGAGGAAGAATAAQAAGAGCAGCAGCACTCACAGCGCCCCAGACTCATGACTCTGAGC	1149				
Ib	122996 CAACTAGCTGGAGTAATTCCACAGAACCAAAGCCCACATCATCTGATCAACCTCG	123055				
OY	1150 gaq	1152				
Lb	123056 AAA	123058				
RESULT 12						
CHECK.						
LOCUS	GELF.L	2232 bp	DNA	INV	04-JUN-1993	
DEFINITION	Caenorhabditis elegans collagen gene.					
ACCESSION	M80650					
VERSION	M80650.1	GI:156249				
KEYWORDS	collagen.					
SOURCE	Caenorhabditis elegans (strain NZ).					
ORGANISM	Caenorhabditis elegans Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdillia; Rhabditoidea; Rhabdillidae; Peloderinae; Caenorhabditis. 1 (sites)					
REFERENCE	Cox,G.N., Fields,C.A., Kramer,J.M., Rossonowicz,H. and Hirsch,D. Sequence comparisons of developmentally regulated collagen genes of Caenorhabditis elegans Gene 76, 331-344 (1989) Journal MEDLINE 893261x1					
AUTHORS	Hird,D.M.					
TITLE	Sequence comparison of the Caenorhabditis elegans dpy-13 and col-34 genes, and their deduced collagen products					
JOURNAL MEDLINE	gene 120, 261-266 (1992) 93013043					
FEATURES						
SOURCE	location/Qualifiers 1..2232 organism="Caenorhabditis elegans" strain="NZ" /db_xref="taxon:6239" 381..393 533..539 /gene-"alpha-collagen" join(585..665,720..1362,1411..1583) /gene-"alpha-collagen" /codon_start=1 /product-"alpha-collagen" /protein_id="AAA27985.1" /db_xref="GI:156250" /translatioin="MODELTRKAVRFVAFSAVFSSVAVLSCVTLEPVRYYVMHRR TMNEITFCRKASAKDINVEVALKSTENSNRTAQVNDAAVTSGGSQNSCESTCLE GPDPACPCKGRPKRGKGALGNNGNRPQQGLPTTPWKKCPQGPGIPGPPG PYDSGEPSGPLPCGDIAADPEPRKPGRKPYPAPGAEHEQSDFRGDELIDGGDG PIFGAGPGSPGPSGPCGAGATCSHPQGRKNIGNDPOFGAGNNICATNPACPNOSKHGH RCETPKRYCALDGVGFEDGTR" 666..7719					
INITIATION						

[illegible]


```

CDS
      join(31073..31155,31211..31304,31350..31526)
      /gene="F36A4.3"
      /codon_start=1
      /evidence-not-experimental
      /protein_id="AA096162.1"
      /db_xref="GI:1255809"
      /translation="MEFEELDINPEGALWCELYLEDVVPPILLRIDRFTSKSK
      MFRIVPGSDPEYVDELYTKIININ;SADGETYQVPRKSKVNGRGSVDFDIDA
      KPHCKPSCKKPPYF"
      32592..33126
      /gene="F36A4.2"
      /join(32592..32728,32795..32885,32932..33126)
      /gene="F36A4.2"
      /evidence-not-experimental
      /protein_id="AA096163.1"
      /db_xref="GI:1255810"
      /translation="MILEPTLVILCSLSQVLFQAKVTNIGTYGELYLEED
      LEHDLADKFTSDMRKQLYTNPKPEPTMDEITYKKNIN;TADGRTYQVPRK
      KKHVWGHREVKELIDASNGKPKTEMPFFYGFWS"

Query Match      3.38; Score 38.6; DH 30; Length 42936;
Post local similarity 49.38; Pred. No. 3.7;
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 772 gatcgcacattacgaagcaacacatcgaatgatatgaaacacatgaatatacctga 831
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5200 GAGCCCAAGAGACACCAAGATCCCAAGACCAAGACCAAGACCAAGATCCCAAGACCA 5141

QY 832 gatgaagaaattggtgacataatttattcttgatcgcgaagcgcacccctcccaaaa 891
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5140 AAGAGATGTTAAGAGAGATATTATTAAACCTGTTCAAGAGACCAAGACCAAGACCA 5081

QY 892 ctggaatggagcaaacacctcgtggtgacatttggaacacccctgtgacctgacgaag 951
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5080 GAGCAATATGAGATGACATGCTCAGCCAGAGAGAGATGAGAACCTTGAGAGACCAAG 5021

QY 952 gacgcacaaacacacatgacgaagaaa 976
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 5020 CGGCAACCAAGACCAAGATCCCAAGACCA 4996

RESULT 14
CEC24F3
DEFINITION
  Caenorhabditis elegans cosmid C24F3, complete sequence.
ACCESSION
  AL022716.1 GI:3642002
VERSION
  AL022716.1 GI:3642002
KEYWORDS
  HTG.
SOURCE
  Caenorhabditis elegans.
ORGANISM
  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
  Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
  1 (bases 1 to 14544)
REFERENCE
  1
  TITLE
    none.
  JOURNAL
    Genome sequence of the nematode C. elegans: a platform for
    investigating biology. The C. elegans Sequencing Consortium
    Science 282 (5396), 2012-2018 (1998)
  99069613
  REMARK
    The C. elegans Sequencing Consortium.
    Erratum: [published errata appear in Science 1999 Jan
    1:284(5398):35 and 1999 Mar 26:284(5410):2103 and 1999 Sep
    3:285(5433):14931]
    3:285(5433):14931
    2 (bases 1 to 14544)
REFERENCE
  AUTHORS
    McManis, A.A.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (27-Apr-1998) Nematode Sequencing Project, Sanger Centre,
    Hinxton, Cambridgeshire CB10 1HQ, England and Department of Genetics,
    Washington University, St. Louis, MO 63110, USA. E-mail:
    jcs@sanger.ac.uk or rwnematode.wustl.edu
  on Sep 22, 1998 this sequence version replaced gi:3093277.
  Coding sequences below are predicted from computer analysis, using
  predictions from Genolinder (P. Green, U. Washington), and other

```

```

available information.
For a graphical representation of this sequence and its analysis
see:
  http://webc.sanger.ac.uk/cgi-
  bin/display2db-wormaccsclass-Sequence%20bject-C24F3
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: this sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
IMPORTANT: this sequence is not the entire insert of clone C24F3.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone C24F3 is at 1 in this sequence. The true
left end of clone F01010 is at 14441 in this sequence. The start of
this sequence (1..104) overlaps with the end of sequence AL021770
the end of this sequence (14441..14544) overlaps with the start of
sequence Z81055.
Location/Qualifiers
  1..14544
  /organism="Caenorhabditis elegans"
  /db_xref="taxon:6239"
  /chromosome="IV"
  /clone="C24F3"
  complement(join(3350..3553,3600..4079,4128..4203,
  4336..4433,5787..5857,5907..6021,6071..6159,6204..6472,
  6420..6764,6810..6945,6997..7231,7487..7580))
  /gene="C24F3.4"
  complement(join(3350..3553,3600..4079,4128..4203,
  4336..4433,5787..5857,5907..6021,6071..6159,6204..6472,
  6420..6764,6810..6945,6997..7231,7487..7580))
  /gene="C24F3.4"
  /note="cDNA EST CEESW54F comes from this gene"
  cDNA EST yk8996.3 comes from this gene
  cDNA EST yk2594.3 comes from this gene
  cDNA EST yk36204.3 comes from this gene
  cDNA EST yk850242.3 comes from this gene
  cDNA EST yk8996.5 comes from this gene
  cDNA EST yk641a2.3 comes from this gene
  cDNA EST yk2594.5 comes from this gene
  cDNA EST yk36204.5 comes from this gene
  cDNA EST yk850242.5 comes from this gene
  cDNA EST yk641a2.5 comes from this gene"
  /codon_start=1
  /protein_id="CAA18773.1"
  /db_xref="GI:3874484"
  /db_xref="SPT:RMBH:093XK6"
  /translation="MNWDRKRVATCTVNNMAIDPKNVEIKVCEAAALZAKIR
  LPELEIPRYGCAHDFELTERHSEMSKLEKSKWPNILVYGLTFRGLLYN
  CAALRNKRLFTIRAKMGLADNVNREKSEKWEKTEKHVOMPLNSDITPQETVF
  GGLTESSNVRLEIGFEICELMSASTVRLAEQGVDMCNCS3SHILGKNMYLNO
  LILSSAVGVAVYLANRGCGDGVVYDAGSSAVNONDILAOIHQPH11ED SVYSAV
  VILSDNCRHRKSSSDKGNASIVVAVLPRIQKMTGGIKNKSFAIPBNVDIOLIS
  PLALCHGPATVLYVLRKSCAGYVLPISGQDSAAVAAMVIMFVAGCAIKRKE
  TGDSDDPAYVLAGKKVGEDPDLQNVLETCYMASEHSDETRQCAEILAKVNSHC
  IFGIDTIVTSILIKVNAVYGEFSPSPENRETMIONQATRAVLSTYFQULVS
  HKRPGGLVLGTANDESLEVLTETKYSNAINLIGSKKDLQPLETAYEKYGA
  ALREVLDSPITPAIKLVNKGAKVDIATGATITDYSISVIGLRKPGTNGYQMTLKI
  LQMGKKSITDTERKVKKFTWRYVNRKRAVSDPAIHAENVSDPHDRKPLTYP
  DTSYOFERIKREVKKNS1"
  join(4753..4834,4885..5678)
  /gene="C24F3.6"
  join(4753..4834,4885..5678)
  /gene="C24F3.6"
  /note="Predicted using genewise
  similar to CULLIN COLLAGN1
  cDNA EST yk9244.5 comes from this gene
  cDNA EST yk9244.3 comes from this gene"

```


* consists of 98 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	658: contig of 658 bp in length	25991	26855: contig of 865 bp in length
659	gap of unknown length	26856	gap of unknown length
1295	contig of 647 bp in length	27764	gap of unknown length
2007	gap of unknown length	28489	gap of unknown length
2721	contig of 714 bp in length	29172	gap of unknown length
3459	contig of 648 bp in length	29923	gap of unknown length
4086	contig of 727 bp in length	30634	contig of 712 bp in length
4801	contig of 715 bp in length	31461	gap of unknown length
5446	contig of 645 bp in length	31462	gap of unknown length
6376	contig of 930 bp in length	32124	gap of unknown length
6747	contig of 471 bp in length	32983	gap of unknown length
7373	contig of 626 bp in length	33747	gap of unknown length
7989	contig of 616 bp in length	34504	gap of unknown length
9108	contig of 1119 bp in length	35273	gap of unknown length
9727	contig of 619 bp in length	36366	gap of unknown length
10439	contig of 712 bp in length	37458	contig of 1092 bp in length
11447	contig of 1008 bp in length	37459	gap of unknown length
12247	contig of 800 bp in length	38610	gap of unknown length
13091	contig of 844 bp in length	39416	gap of unknown length
13788	contig of 697 bp in length	40019	contig of 703 bp in length
14602	contig of 814 bp in length	41351	gap of unknown length
15418	contig of 816 bp in length	41352	gap of unknown length
15966	contig of 548 bp in length	43046	contig of 1694 bp in length
16948	contig of 982 bp in length	43892	contig of 847 bp in length
17661	contig of 713 bp in length	44893	gap of unknown length
18453	contig of 792 bp in length	45469	contig of 1577 bp in length
19243	contig of 790 bp in length	45470	gap of unknown length
19998	contig of 755 bp in length	46763	contig of 1294 bp in length
20862	contig of 864 bp in length	47850	contig of 1088 bp in length
22034	contig of 1172 bp in length	47851	gap of unknown length
23296	contig of 1262 bp in length	49437	contig of 1587 bp in length
24124	contig of 828 bp in length	50702	contig of 1265 bp in length
25153	contig of 1029 bp in length	51982	gap of unknown length
25990	contig of 847 bp in length	51983	gap of unknown length
		54151	contig of 1169 bp in length
		54352	contig of 842 bp in length
		54994	gap of unknown length
		55057	contig of 1064 bp in length
		56117	gap of unknown length
		57311	gap of unknown length
		57312	gap of unknown length
		58302	contig of 991 bp in length
		58403	gap of unknown length
		59260	contig of 958 bp in length
		59261	gap of unknown length
		60675	contig of 1415 bp in length
		60676	gap of unknown length
		61873	contig of 1198 bp in length
		61874	gap of unknown length
		63206	contig of 1333 bp in length
		64718	gap of unknown length
			contig of 1512 bp in length

Genome version 4.5
Copyright (c) 1994 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 1, 2001, 22:43:50 : Search time 75.08 seconds
(without alignments)
430,487 Million cell updates/sec

Title: US-09-645-321-1

Percent score: 2147
Sequence: 1 MIAHRQELADQYYQALHDP.....CFIVPRALMLCAETLQPV 391

Scoring table:
Gap: 10.0, Export: 0.5

Search: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match: 0%
Maximum Match: 100%

Listing first 45 summaries

Database: PIR-65:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Printed. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2147	100.0	391	2 S75649	reinin-binding prot
2	694	32.3	402	2 A35741	reinin-binding prot
3	668	41.1	419	2 JX0187	reinin-binding prot
4	645	30.0	417	2 JX0188	reinin-binding prot
5	148	6.3	418	2 S40824	hypothetical 48k p
6	108.5	5.1	607	2 S77092	hypothetical prote
7	103	4.8	662	2 S55387	hemocyanin precurs
8	98.5	4.5	689	2 S66006	conserved hypotet
9	97.5	4.5	1007	2 T42219	alpha-mannosidase
10	95	4.3	876	2 G70864	probable val's prot
11	94.5	4.3	266	2 T20806	hypothetical prote
12	94.5	4.4	2493	2 A55481	adenylate cyclase
13	93	4.3	657	1 HBU08	hemocyanin chain b
14	92.5	4.3	748	2 S19652	cellulocyclinase c
15	92	4.3	1067	2 H75139	isolouy1 tRNA syn
16	91.5	4.3	692	2 B70484	conserved hypotet
17	91	4.2	391	2 T10239	hypothetical prote
18	91	4.2	657	1 BHDVA	hemocyanin chain a
19	91	4.2	961	2 JRC200	alpha-mannosidase
20	90	4.2	417	1 S49051	cytochrome p450 ty
21	89.5	4.2	583	2 JG6504	dephar, alpha-trical
22	88.5	4.1	698	2 S52674	general sporulat
23	88.5	4.1	900	2 T39530	hypothetical prote
24	88	4.1	900	2 S73748	alanine-tRNA lig
25	87	4.1	348	2 H75409	glycosyl hydrolase
26	87	4.1	720	2 JG6562	heat shock protein
27	86.5	4.0	642	2 H69466	conserved hypotet
28	86	4.0	1066	2 F71100	isolouy1 tRNA 1
29	85.5	4.0	316	2 T26766	hypothetical prote

30	85.5	4.0	686	1 ALDYAT	amylase A (EC 3.2
31	85.5	4.0	912	2 D72644	hypothetical prote
32	85.5	4.0	2405	2 T08164	dynactin alpha heavy
33	85	4.0	359	1 M0H9CD	N-acetylaminoglyc
34	85	4.0	700	2 S57194	calpain (p3.4.22
35	85	4.0	833	2 S62136	cd44 protein (p3.4
36	84.5	3.9	180	2 A32246	70k antiapn - chla
37	84.5	3.9	536	2 A81710	ribosomal protein
38	84.5	3.9	580	2 D72093	ribosomal protein
39	84.5	3.9	748	2 S41050	fibroblast growth
40	84.5	3.9	750	2 S41051	fibroblast growth
41	84	3.9	465	2 D69785	beta-glucosidase b
42	84	3.9	2222	1 A36028	DNA-directed RNA p
43	83.5	3.9	454	2 G64737	polynucleotide ade
44	83.5	3.9	470	1 NM1VW8	exo-alpha-stalidas
45	83.5	3.9	513	2 S48981	hypothetical prote

ALIGNMENTS

RESULT 1
S75649
reinin-binding protein-related protein - Synchocystis sp. (strain PCC 6803)
N:Alternate names: protein slr1975
C:Species: Synchocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25 Apr 1997 #text change 20-Jun-2000
C:Accession: S75649
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K. Okumura, S.; Shimpo, S.; Takeuchi, G.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu, K. DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp.
S:
A:Reference number: S743322; M01D:97061201
A:Accession: S75649
A>Status: nucleic acid sequence not shown; Translation not shown
A:Molecule type: DNA
A:Residues: 1-391; KAN.
A:Cross-references: EMBL:090912; GIB:AB001339; NID:01653228; PIR:66A32101; U00019
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996.

Query Match 100.0% Score 2147; 108.2; Length 391;
Best local similarity 100.0%; Pred. No. 2; 16-176;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIAHRQELADQYYQALHDPVLPFEWERYSLDKSGSYETGLDKRGVETDKEFTWQNRQ 60
DB 1 MIAHRQELADQYYQALHDPVLPFEWERYSLDKSGSYETGLDKRGVETDKEFTWQNRQ 60

QY 61 VMOFAVYVNRKLEPRQWLEIARRGADPLAAGRCRQYQNNYFALDQGRKPLROYNNVSYG 120
DB 61 VMOFAVYVNRKLEPRQWLEIARRGADPLAAGRCRQYQNNYFALDQGRKPLROYNNVSYG 120

QY 121 FAAMASQYVALASQACAKAIALQAYNNVLRQHNKRGQYKSYSPGIRPLKSLAVNPILA 180
DB 121 FAAMASQYVALASQACAKAIALQAYNNVLRQHNKRGQYKSYSPGIRPLKSLAVNPILA 180

QY 181 NLTLEMWLLPRTIVEEVLADQVREVMTDLPDPIGIMREAVPTSGEYVSFEKRLNPG 240
DB 181 NLTLEMWLLPRTIVEEVLADQVREVMTDLPDPIGIMREAVPTSGEYVSFEKRLNPG 240

QY 241 HGTEAMFPMHLAGKSLKRLQGLQALAVVNIITFYADDERFGGIPYPLDQGHVQVQIEM 300
DB 241 HGTEAMFPMHLAGKSLKRLQGLQALAVVNIITFYADDERFGGIPYPLDQGHVQVQIEM 300

QY 301 DQKLMWHLETLVALAKGHDATGOEKQWQFERKSHYAMSNIPDERGEMFGYLNRGGEV 360
DB 301 DQKLMWHLETLVALAKGHDATGOEKQWQFERKSHYAMSNIPDERGEMFGYLNRGGEV 360

QY 361 LNLKAAKMKRCFIVPRALMLCAETLQPV 391
DB 361 LNLKAAKMKRCFIVPRALMLCAETLQPV 391

62 WQFVNTNRELE -PKQWLEIARIADFLARIR- IQDNNWTFALDQEKPLQRYNVE 117
 118 SPTFAAMAFSYALASQJAKAKATALQAYNNVIR-RQHNRKQYERKSYQIRPKSLAVP 176
 124 SEFTYIMAMNLMKQVTFVRQTEAVEEMQQLVHWVEQDASGRLPOLQCAFAAEFRAVP 182
 177 MLIANI LHEMLILPPIVEVAQIVRYV-----MIDPDEIDIMKEAVIPDEEV 229
 184 MMLINVEQI-----GPAPEELAGKAYAEIGDQWARRILQHVQDQDAVLENSDQKEL 236
 240 TSPEDLLNPHGIEAMVPMKIDAJRSIDRLQLEQAI-AVNLITLEYAMTEERSGITFYL 288
 247 PQTJRKQVPHLEACQPFILRHCRKQDPELKAHVILKEKILLPESGMDPDHGLTFYQ 296
 289 DRGHPQVQITWQKILWVHLELVALAKQHQALGQPKWQWPHFRVDMYAKSHADPEYG 348
 297 LAINPFTQLEFAMKRLWMPHSEAMIAFLMYSISGDFVILRLYQVAYETFEQDFDEYG 356
 449 FWPYENRHRDVLINIKGKKKQCEHVIRALWIAFELL 386
 457 FWPYENRHRDVALSTKGKPKQCEHVIRALWIAFELL 394
 RESULT 5
 S40824
 Hypothetical 48k protein (qlna-fdhe region) - Escherichia coli
 Naitomate names: hypothetical protein t418
 CDate: 19-May-1994 #sequence_revision 01-Sep-1995 #text_change 04-Mar-2000
 CAccession: S40824; #6193
 RefLink: 111, G.; Burland, V.; Daniels, D.L.; Hattner, F.K.
 Nucleic Acids Res. 21, 3391-3398, 1993
 ATitle: Analysis of the Escherichia coli genome. 111. DNA sequence of the region from 8
 AReference number: S40802; M01D:93347969
 AAccession: S40824
 AStatus: preliminary; nucleic acid sequence not shown; translation not shown
 AModel type: DNA
 AResidues: 1-418 - PUA
 ACross-references: EMBL:J19201; NID:q404961; PUD:AA03013.1; PID:q304984
 ANote: The nucleotide sequence was submitted to the EMBL Data Library, October 1993
 RefLink: F.K.; Plunkett, 111, G.; Bloch, G.A.; Fernu, N.T.; Burland, V.; Riley, M.; G.
 A. Rose, D.J.; May, R.; Shao, Y.
 Science 272, 1454-1457, 1997
 ATitle: The complete genome sequence of Escherichia coli K-12.
 AReference number: A64720; M01D:97426617
 AAccession: 66193
 AStatus: preliminary; nucleic acid sequence not shown; translation not shown
 AModel type: DNA
 AResidues: 1-418 - BLAL
 ACross-references: GBA:AB000464; GBA:000096; NID:q2467324; PUD:AA013442.1; PID:q1790313
 AExperimental source: strain K-12, substrain MG1655
 AGenes:
 ANotes:
 ASuperfamily: Escherichia coli hypothetical 48k protein (qlna-fdhe region)
 Query Match 6.98; Score 148; Dh 2; Length 418;
 Best Local Similarity 21.14; Pred. No. 5.5e-05;
 Matches 82; Conservative 63; Mismatches 195; Indels 78; Gaps 17;
 48 PQTJRKQV FTHKFLMIONROVQFAVYNLEIKPQWLEIARHGAFLARHG----- 92
 14 PQTJRKQV FTHKFLMIONROVQFAVYNLEIKPQWLEIARHGAFLARHG----- 92
 38 PQTJRKQV FTHKFLMIONROVQFAVYNLEIKPQWLEIARHGAFLARHG----- 92
 93 -----RQGFQNVYFALLQEG--KPLQFYNYVSDTFAMAFSYALASQJAKAKAI----- 141
 88 NGALRKKVYLAWAVVNDQSVADSKQYQHF---FALLQAS-AVNLITLEYAMTEERSGITFYL 143
 142 -----ALQATYNNVIRKQHNKQYERKSYQIRPKSLAVNMLIANILHEM 186
 144 IELIKYVWSEEDQLESDQESKTEYRG-----NNANMLAVEFLIYVDTIRK 196

187 EMLPPTVEEVIQVIRE-----VMIDFLPELQMKR- AVTPGEEVDSFSLNNG 240
 197 KMLDRAIIVASVILHVANNHRYVNEHPTQMNPIIYQNKINIAHRR-KAQG---LHG 252
 241 DSIETAMTMDIAQRSGDR-----QLQDAIAVNLITLEYAM-DEERSGITFYLROGH 293
 253 HWLEWGRMLHIDALEAKQEQFPAWLEDAKTEFATKRAVADGAGAVYTIWEEK 312
 294 PPTQLEWQKILWVHLELVALAKQHQALGQPKWQWPHFRVDMYAKSHADPEYG 353
 314 PVVR -EKVWPIVEAMGTAVALTITGQYETIWTWMEY-ITKIMQENSMQJE 368
 354 INRGVLENLKSGKKR - GGFHVR 377
 359 LIAQNKVILKVMGKQDLYHILHCTVIR 397

RESULT 6
 S77092
 Hypothetical protein slr185c - Synechocystis sp. (strain pcc 6803)
 CSpecies: Synechocystis sp.
 AVarities: pcc 6803
 CDate: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 CAccession: S77092
 RefLink: T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimizu, S.; Iakouchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-116, 1996
 ATitle: Sequence analysis of the genome of the unicellular cyanobacterium Syn. 6
 AReference number: S74322; M01D:97061201
 AAccession: S77092
 AStatus: nucleic acid sequence not shown; translation not shown
 AModel type: DNA
 AResidues: 1-607 - KAN-
 ACross-references: EMBL:D90908; GBA:AB001339; NID:q1652725; PUD:BA017650.1; PID:q165
 ANote: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
 CSuperfamily: Synechocystis hypothetical protein slr185c

Query Match 5.18; Score 108.5; Dh 2; Length 607;
 Best Local Similarity 22.78; Pred. No. 0.22;
 Matches 89; Conservative 47; Mismatches 127; Indels 129; Gaps 25;
 58 NROVQFAVYNLEIKPQWLEIARHGAFLARHG--IQDNNWTFALDQEKPLQRYNVE 112
 159 SKLVYGRATIAF-MITGDQVITAEKQGTETIRHMRPVKDEDIYWHATDVQGRKPK 217
 113 PY-NVSTQFAM-AFSQ-YALASQJAKAKATALQAYNNVLRQ--HN----- 155
 218 FASRGQDYDAIPAYHQLYALAG-----PIQIYKCIQIRKILHAKQITIKIFDKFF 269
 156 -IKQYERKSYQIRPKL-----KSLA-----VMILIANIHEMLILPPI 194
 270 LKSEGVYFSLDPLMLDPRESLSPNKAQKMNNSVGDHAYLNL-----WL--AT 321
 195 VEVVIA-----QIVKRWIIDLPIDIGIMRAVPIQGFV-----DSFHQRL 337
 322 GQKVAIDMLTYTITIKTYVDT-----DHSFQDERHYEMASHDITWQWQ 368
 238 N--PMSITFAMFMKIDAJRSIDRLQLEQAI-AVNLITLEYAMTEERSGITFYLDR--QG 292
 369 NNAVVGINKLAIANNIMHNSLKKSEKYYVHAKKIALIMAVQSDQKQWYIVVERILIN 428
 293 HPI-QQLEW-DQKILWVHLELVA-----LA-----KQDAIQLQKQWQWPHFRVDMYAKSH 342
 429 HSRCHQFVWIRKQVADQEQALIVLILMILDEEYHRRQE----- 471
 343 ADPEYGHVQVYNNRQ--FVILN-----LKG 367
 472 ASAFYNAWPIQIDFQKQYFNVANCIPIYLAG 503

QY 56 LQNKVWQFAVYNNKLEKPKQM-----LELAHQADPLARHQKQ- DANNYF----- 101
 DB 120 VE-----LAFSTR-----MMHQUTNAICEVVD-----LVKQGLEFAAGGVMNNEA 162
 QY 102 -----ALNCTKPLKQYNNF--SQTFAAMAFSOYALASQVAKALAUAYNNV---RR 152
 DB 163 AHHYATLILNCLLDELPLEDLEKIDRIKRVAMHIDPHQSHKQASLPAQMDPLDFEOL 222
 QY 153 QUNRGOYEKSVPTKRLKSLAVENLLANLTLEMFW-----LPPPTVEVLAQTVKRYM 207
 DB 223 DQVKKRVEE-----NLDG EGVWRASASIKQVPA-----AULFISVL 258
 QY 208 IDPELPELQIMKEAVPTGEVDSFESRIENHQITFAMWEMMDIAQSHDROLQVQAL 267
 DB 259 ENIYNPEKIDWDLCAIKRPV---EQRSEPEYNAEELVNFQALAIADQHPFTNITIM 315
 QY 268 VLNILLEYAMIDEEFGITPYELDKQCHPVQVQIMQKIMWMLFETVALAKG-HQALQDE- 425
 DB 416 IMSDIOFVE-----NANMFRRNDIKLQIVNAQQAQNSRV 351
 QY 326 -----KQWMEFEVDYAMS-----HFAHFEYQGEKRYLNR 357
 DB 452 NLYSTPAVYIWEINKAMIMSVKQDHPYAlphHOBWSYFSSR 497

RESULT 10

probable vols protein - Mycobacterium tuberculosis (strain H37RV)

C/Spect: Mycobacterium tuberculosis

C/Accession: 17-jun-1998 #text_revision 17-jun-1998 #text_change 20-jun-2000

C/Spec: 670864

C/Ref: S.T.: Broese, K.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

C/Ref: Connor, R.; Davies, K.; Devlin, K.; Fellwell, T.; Gellies, S.; Hamilton, N.; Holtroyd, S.

C/Ref: Rajandram, M.A.; Roberts, J.; Rutter, S.; Seeger, K.; Skellern, S.; Squares, S.

C/Ref: Narine, S.; Sultison, J.E.; Taylor, K.; Whitehead, S.; Bartlett, H.G.

C/Ref: A little description of the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MIMD:98295987

A/Accession: 670864

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1 876 670864

A/Cross-references: db:AL021246; db:AL124456; NID:4261507; PIDD:CAA16025.1; PIDD:q279148

A/Experimental source: strain H37RV

C/Genetics:

A/Genes: vals

C/Superfamily: valine--trna ligase

Query Match 4.4% Score 95; DB 2; Length 876;

Best Local Similarity 19.8%; Pred. No. 5.3;

Matches 74; Conservative 47; Mismatches 141; Indels 122; Gaps 18;

QY 25 WEKYSIDQWQVAVPTLQKQVFD--DKEFLMUN-----RQVWQFAVYNNLEP 73
 DB 117 WKES--GATAGQMRKLDGQWSKQRETMEGLSKAVTLEKRYLWLLY-RAER 171
 QY 74 KQMLEIARHQAFLARIRBQDQW-----YFALQDEKPLKQYNNVSTFAMAFSOY 129
 DB 172 LVNWSVIVQIALSLEVYNNKVECHVSRKQSLDLS-----QPHIVATIRVETMID 226
 QY 130 ALASQVAKALAUAYNNVLRQINFRKQYKSEYFGTRPLKSLAVENLLANLTLEMFL 189
 DB 227 ALAV-----HFDERIRHLVGT--SLAHPVDELEAI----- 256
 QY 190 IDPELPELQIMKEAVPTGEVDSFESRIENHQITFAMWEMMDIAQSHDROLQVQAL 267
 DB 257 -----VNEHVDPEFGAVKVI PAID-PIQDFELGVHQLPMPSTIDTK 299
 QY 214 GGLLNKGRITFAMWEMMDIAQSHDROLQVQALAVNLTLEYAMIDEEFGITPYELDKQCH 293
 DB 400 GRLVLDGIRFIRM-----DREFARVAVKQALAAQGVVERKPD-----YLSWQH 344

QY 294 PPOUE-WQDKL---WVHLLEIV---ALAK;- HCATQCKWQMEFRVHVA 448
 DB 445 SERSGEIEERKLSQWVWVESLAKKAGDQVKNLQIVIHASNEPFWFSNVLIMHWWLS 404
 QY 439 ----WSHFAIRFYG 448
 DB 405 KQLWCHRRIPWYG 418

RESULT 11

120806

hypothetical protein F13A7.10 - Caenorhabditis elegans

C/Spec: Caenorhabditis elegans

C/Ref: 15-oct-1999 #sequence_revision 15-oct-1999 #text_change 15-oct-1999

C/Accession: 120806

C/Ref: R.Brady, A.

Submitted to the EMBL data library, March 1997

A/Reference number: Z19427

A/Accession: 120806

A/Status: preliminary; translated from GB/EMBL/DB

A/Residues: 1 266 120806

A/Cross-references: EMBL:Z99477; PIDD:CA807572.1; GSDPB:G000024; CEST:F13A7.10

A/Experimental source: clone F13A7

C/Genetics:

A/Genes: CESP:F13A7.10

A/Map position: 5

A/Inserts: 27/1; 182/1

Query Match 4.4% Score 94.5; DB 2; Length 266;

Best Local Similarity 20.4%; Pred. No. 1.4;

Matches 50; Conservative 33; Mismatches 96; Indels 67; Gaps 9;

QY 109 PLQYVWESICFAAMAFSOYALASQVAKALAUAYNNVLRQINFRKQYKSEYFGTR 168
 DB 34 PSNTPFIVSMRCPDAQKALVAA--VKLIHSEVIN--IMQKPMWYFKNKQGV 87
 QY 169 P-----LKSIAVEMLANLTLEMFLPPTVEEVAQTVKRYLWLLY-RAER 217
 DB 88 PLRHAQCKKIVLSAVIPYIDIPPEK-ILSDPYEKVQCKLID ERLSDQILAFGR 146
 QY 218 MRFAVPTGEVDSFESRIENHQITFAMWEMMDIAQSHDROLQVQALAVNLTLEMFL 277
 DB 147 VTRAIKMPDELKEFS-----ILKATFEA 171
 QY 278 DEEPQGIY-----FIDRQHPPVQLEWQKIMWVH-LEIVAIKQCHQALQDEKW 328
 DB 172 ESILGAFVNSISQGVQVLLVYS-----FQVYWMILFELPPLPSNFRQNGYPKLS 226
 QY 329 QWFERV 344
 DB 227 QWFKAI 242

RESULT 12

A55481

adenylate cyclase (ec:4.6.1.1) nact - smut fungus (Ustilago maydis)

C/Spec: Ustilago maydis (corn smut)

C/Ref: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Sep-1997

C/Accession: A55481

C/Ref: Rold, S.; Duncan, G.; Bartlett, K.; Kronstad, J.

Genes Lev. 8, 2805-2816, 1994

A/Title: CAMP regulates morphogenesis in the fungal pathogen Ustilago maydis

A/Reference number: A55481; MIMD:95087882

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1 2493 602

A/Cross-references: db:U43918; NID:9603939; PIDD:q603940

A/Note: nucleotide sequence not given; amino acid sequence not complete

C/Genetics:


```

07 183 -LEEMWLLPPIVHEVIACTVREVMIDPLEIGIMRAVPIGEVWSP----- 232
10 503 EPYLYIMAEWKGJWEAEV-QAINEVFNMLLIVERISANANTQDTPISVEPHSTIL 561
07 243 -----FOGLINFGHIFAMFM-----MQLAKR-----SGDKQLQHQ 264
10 562 NPNNQJENIYFADNPENIPKQILFSPHLYGDSYFVROQFMDPAQFDCAGLEQDEAAQAR 621
07 265 ALIAYLIN-TELEYAMDEEEFG-----GTFEYELDRGHHPOLEW-----DQKLMV 307
10 622 C-RVIVINIVALEGDMHEHPQYIKHGYGILL-----GTFQGMIMPQAKSSQALRNAN-- 673
07 408 HLEII VALAKHQALGQKQWQ-----WFER--YHDYAMSPADPE----- 346
10 674 -----SHITINVOQMOQAASFEKKGKGINAYWS--MNPESADIMGWLTIPMDP 721
07 447 -----YGRMFGYINRKGIVILNLKG 366
10 722 VIANDMGMQWIGQDPRKLTQILNNMG 747

```

RESULT 15

H75149

Isolated: tRNA synthetase (tles) FAB0616 - *Pyrococcus abyssi* (strain Orsay)
 Species: *Pyrococcus abyssi*
 Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 Accession: H75149
 Reannotation, Genoscope
 Submitted to the EMBL data library: July 1999
 Description: *Pyrococcus abyssi* genome sequence; insights into archaeal chromosome str.
 Reference number: A75001
 Accession: H75149
 Status: preliminary
 Molecule type: tRNA
 Accessions: F-1067-KAW
 Accession: CH:AJ248285; CH:AI096836; NID:95458067; PID:CAH49833.1; PID:9545834
 Experimental source: strain Orsay
 Notes:
 Accession: FAB0616
 Classification: Isolated: tRNA Lysase

Query Match

4.38; Score 92; DB 2; Length 1957;

Post local similarity 20.4%; Pred. No. 13;

Matches 56; Conservative 24; Mismatches 47; Indels 148; Gaps 14;

```

07 153 QHNKGGY-----FKSY-----PCTRPKSLAVYMLIANIL----- 184
10 193 EHEVGGYKIKKQDSIYKFPVEGKENFYLLWITTPWILP-----ANLAVSADP 242
07 185 -----EMWLLPPTVEEVLACTVREVMIDPLEIGIMRAVPIGEVWD 240
10 243 QYDYKVKVVEFGKREYVWILAKAIDKVLG-----ETGV-----KGVVE 282
07 241 SPEGRILNFGHQLFAMWPMMDIAQKSGDKQLQHQALAVVNLIIYAMDEPQALPYFLDR 290
10 281 EFKGREL--EGLRYVHILMD-----EY----- 302
07 291 QHNHGGYFMWQKIMWVH-----LEII VALAKG-----HQALGQKQWQMPKRVHAYAMSHA 344
10 303 -----FQCKFKFKKRYAMHKVLIADPVLLEEGTQIVHIAHQHCP-----E 342
07 344 DEEYEMIVGY-----LNRGEVLLNLKGKRGGE 374
10 343 DEEYDQKTLIVYSPIDQCK-----YIDCKMKGLY 373

```

Search completed: January 1, 2001, 22:54:19
 Job time: 1829 sec

